DUKE MATHEMATICAL JOURNAL

EDITED BY

ARTHUR BYRON COBLE DAVID VERNON WIDDER

JOSEPH MILLER THOMAS
Managing Editor

VOLUME II
1936

DUKE UNIVERSITY PRESS
DURHAM, N. C.

CONTENTS

VOLUME 2, 1936

GROUPS OF CREMONA TRANSFORMATIONS IN SPACE OF PLANAR TYPE

By Arthur B. Coble

1. Introduction. We shall say that a group G of space Cremona transformations is *of planar type* if it possesses the distinguishing characteristics of the entire group of Cremona transformations in the plane. The characteristics which we shall stress are the following:

(α) G has an infinite continuous set of generators all of the same type.

In the plane this set is the set of quadratic transformations with distinct F-points.

(β) A particular element of G is defined by the choice of certain discontinuous parameters, positive or zero integers, which fix the *type* of the element (i.e., the nature of its F-system), and of certain continuous parameters which fix the position of its F-system.

This requirement rules out the group of inversions in space which has only two types of elements, namely: the collineation, and the quadratic transformation with a simple F-point and with a conic as an F-curve of the first kind.

(γ) Associated with G there is a group g of linear transformations on an unrestricted number of variables with integer coefficients. Each element of g defines a type of element in G. The product of two elements of G has a type defined by the product of the corresponding elements in g.

(δ) The linear group g of types in G has a linear and a quadratic invariant.

The number of groups of the type indicated which have thus far been exhibited is quite limited. In each space S_n ($n \geqq 2$) there is the group of "regular Cremona transformations",[1] which has interesting applications. These transformations have been called "punctual" by Miss Hudson.[2]

In S_3 there is a group whose generators are the cubic transformations which have a degenerate sextic F-curve of the first kind made up of a space cubic curve, fixed for the entire group, and of three variable bisecants of this curve. Montesano[3] has shown that in this group the types are isomorphic with the ternary types.

Snyder[4] reports a somewhat more special type of cubic transformation whose

Received December 20, 1935.

[1] A. B. Coble, *Point sets and allied Cremona groups* II, Transactions of the American Mathematical Society, vol. 17 (1916), pp. 345–385.

[2] Hilda P. Hudson, *Cremona Transformations in Plane and Space*, Cambridge University Press, 1927.

[3] D. Montesano, *Su alcuni tipi di corrispondenze cremoniane spaziali collegati alle corrispondenze birazionali piane di ordine n*, Napoli Rendiconti, (3), vol. 27 (1921), pp. 164–175.

[4] V. Snyder, *Some recent contributions to algebraic geometry*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 673–687.

sextic F-curve breaks up into four skew lines and their two transversals. With the two transversals fixed (for convenience) this type serves as a generator of a group of ternary type.

In this paper we develop a novel group G of space transformations whose elements have, in addition to isolated F-points, a fixed F-curve C of the first kind which is a generic space sextic of genus four. The elements are therefore definitely not products of cubic transformations. The generating type of element for this group is obtained in **§2**. It is a de Jonquières or monoidal involution of order four with one isolated F-point. This generating involution is listed by Sharpe and Snyder [[5]**§5, §10**] with only the briefest mention. Since it is fundamental for the group G we develop it here somewhat more fully.

The type of homaloidal web for the generic element of G is obtained in **§3**. This web depends upon a certain integer "characteristic", and the linear group g of this characteristic with a quadratic and linear invariant is discussed in **§4**. In **§§5, 6** the relation between corresponding elements of g and of G, and the effect of these elements upon the characteristic of a linear system, are discussed. In a later paper a number of other groups of this planar type will be given. The one discussed here is of exceptionally simple character.

2. The Cremona involution I_p of order four. Let C be a generic space sextic curve of genus 4 on a quadric Q. If then C is the complete intersection of Q with a cubic surface K, the system (∞^4) of cubic surfaces on C has the form

$$(1) \qquad \alpha_0 K + (\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4)\, Q \equiv \alpha_0 K + \pi Q = 0.$$

The generic pencil of cubic surfaces on C contains one member of the form πQ. The base curve of the pencil is the common curve of some K and πQ. It is therefore the sextic curve (K, Q) and the plane cubic curve (K, π), these two curves meeting in six points (K, Q, π) which are on the conic (π, Q) on π. The generic net of cubic surfaces on C contains a pencil of the form πQ and is obtained by adding to the above pencil another surface, $\pi' Q$. Since Q meets the residual cubic (K, π) in the six points mentioned on C, and π' meets this cubic in three points on the line (π, π'), we see that

(2) *Three generic cubic surfaces on the sextic curve C meet outside of C in three collinear points.*

The system (∞^3), or *web*, of cubic surfaces on C and a point P not on Q is also of the form (1), $\alpha_0 K + \pi Q = 0$. There follows at once that

(3) *The web of cubic surfaces on C and on a fixed point P not on Q is such that the net of surfaces of the web on a point x is on a second point x', x and x' being collinear with P. The points x, x' are partners in a Cremona involution I_P, perspective with center at P.*

It is easy to set up equations for the involution and the analytic method has certain advantages over the synthetic method. Let P be the point 1, 0, 0, 0 and

[5] F. R. Sharpe and V. Snyder, *Certain types of involutorial space transformations*, Transactions of the American Mathematical Society, vol. 21 (1920), pp. 52–78.

let K be that cubic surface on C which has a node at P. Since P is not on Q, we can take $x_1 = 0$ to be the polar plane of P as to Q. Then K and Q have the form

$$(4)\qquad \begin{aligned} K &\equiv k_2(x_2x_3x_4)x_1 + k_3(x_2x_3x_4) = 0, \\ Q &\equiv x_1^2 + l_2(x_2x_3x_4) = 0, \end{aligned}$$

where k_2, k_3, l_2 are ternary forms of orders 2, 3, 2 respectively in x_2, x_3, x_4. The web of cubic surfaces on P has the form

$$(5)\qquad \alpha_0 K + (\alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4)Q = 0.$$

The condition on the pair x, x' is that they yield in (5) the same condition on α_0, α_2, α_3, α_4. Thus

$$(6)\qquad \left\| \begin{matrix} K(x) & x_2Q(x) & x_3Q(x) & x_4Q(x) \\ K(x') & x_2'Q(x') & x_3'Q(x') & x_4'Q(x') \end{matrix} \right\| = 0.$$

Since x, x' are collinear with P, we may take

$$x_1' = x_1 + \lambda, \qquad x_2' = x_2, \qquad x_3' = x_3, \qquad x_4' = x_4.$$

The conditions (6) are then satisfied if we determine λ so that $K(x)\,Q(x') \equiv K(x')\,Q(x)$. Replacing the values of x' and using the forms (4), we find that

$$\lambda = (-k_2x_1^2 - 2k_3x_1 + k_2l_2)/(k_2x_1 + k_3),$$

$$x_1 + \lambda = (-k_3x_1 + k_2l_2)/(k_2x_1 + k_3).$$

Hence the equations of I_P are

$$(7)\qquad x_1' = -k_3x_1 + k_2l_2, \qquad x_i' = x_i(k_2x_1 + k_3) \qquad (i = 2, 3, 4).$$

The locus of fixed points of I_P for which $x_i'/x_i = x_j'/x_j$ is

$$(8)\qquad F = k_2x_1^2 + 2k_3x_1 - k_2l_2 = 0.$$

This is a quartic surface with node at P and the polar cubic surface of P is K. Eliminating x_1 from K and F we get $k_2(k_3^2 + k_2^2l_2) = 0$. The factor

$$(9)\qquad G = k_3^2 + k_2^2l_2 = 0$$

is the sextic cone from P to C as may be seen by eliminating x_1 from (4). The factor k_2 indicates the six lines

$$(10)\qquad F_2\colon \quad k_2 = k_3 = 0,$$

which are on P and bisecant to C. Hence

(11) *If F is the quartic surface, with the same node P and nodal tangents as K, which touches the cone G from P to C along C, then the quartic involution I_P is the locus of pairs x, x' on a line λ through P and harmonic to the two further points in which λ on the node P of F meets F again.*

It is clear from the equations (7) of I_P that K is the P-surface which cor-

responds to the isolated F-point P. It is also clear from (11) that C is an F-curve of the first kind whose P-surface is the sextic cone G. Hence

(12) *The homaloidal web* (10) *is the system* $(P^3C)^4$. *The point P is an isolated F-point whose P-surface is* $K = (P^2C)^3$. *The F-curves of the second kind are the six lines F_2 on P and bisecant to C.*

That we have enumerated all the isolated F-points and F-curves of the first kind is clear since the web $(P^3C)^4$ is transformed by I_P into $(P^{12}C^4)^{16} = \{(P^2C)^3\}^3 \cdot (P^6C)^6 \cdot (P^0C^0)^1$, i.e., into the web of planes. Also two members of the homaloidal system $(P^3C)^4$ meet in a curve of order 16 with 9-fold point at P. From this, C of order 6 and F_2 of order 6 and 6-fold at P separate, leaving a quartic curve with triple point at P, which is the proper transform by I_P of a line. We have therefore all the F-curves of the second kind.

A point on Q has coördinates $\sqrt{-l_2}, x_2, x_3, x_4$. This yields in (7) the point $x' = -\sqrt{-l_2}, x_2, x_3, x_4$. Hence Q is invariant under I_P. The definition (11) of I_P may therefore be replaced by the following which makes no use of F:

(13) *A line λ on P meets K and Q each in two pairs of distinct points. A pair x, x' of I_P on λ is a pair of the involution defined by the two pairs mentioned.*

3. The Cremona group G generated by involutions I_P. Let $P_1, P_2, \cdots$ be any sequence of points in S_3 which are not on Q. Then the sequence of involutions, $I_{P_1}, I_{P_2}, \cdots$, has a definite product which is by definition a particular element of a Cremona group G. We ordinarily assume in forming such a product that the point P_k is in generic position with respect to the F-points of $I_{P_1} I_{P_2} \cdots I_{P_{k-1}}$, or else that it coincides with one of them. If P_k were, for example, on a P-surface of the product just mentioned, we would have a case of coalescent ordinary singularities.

An examination of a few of the simpler products indicates the following theorem:

(1) *The generic element of the group G has a homaloidal web of the form*

$$(C^{x_0-1} p_1^{3x_1} p_2^{3x_2} \cdots p_k^{3x_k})^{3x_0-2},$$

where $x_1, \cdots, x_k$ have zero or positive values and

$$L \equiv x_0 - x_1 - x_2 - \cdots - x_k = 1.$$

It is to be observed that the isolated F-points $p_1, p_2, \cdots$ are not the F-points $P_1, P_2, \cdots$ of the involutions which generate the above element. Thus under $I_{P_1} I_{P_2}$ planes pass into members of a web with F-points at $p_2 = P_2$ and $p_1 =$ transform of P_1 by I_{P_2}.

We note that the theorem is true for $x_0 = 1$ and $x_i = 0$ $(i > 0)$. The web in this case is the web of planes and the element of G is the identical collineation since we wish to keep C fixed and C admits no other collineations. For $x_0 = 2$, $x_1 = 1, x_i = 0$ $(i > 1)$, the web is that of I_{P_1} in **2**(12). To prove the theorem it is therefore necessary only to show that the web (1) goes into a similarly defined web under I_{p_1} and under $I_{p_{k+1}}$. Under I_{p_1} the position of the F-points $p_2, \cdots, p_k$ will change, and under $I_{p_{k+1}}$ the positions of $p_1, \cdots, p_k$ will change. If, how-

ever, we assume the multiplicity $x_{k+1} = 0$ at p_{k+1} for the given web, then, under either involution, we find that the change in the multiplicities x is given by the equations

$$\text{(2)} \qquad i_j: \quad \begin{aligned} x_0' &= 2x_0 - 3x_j, \\ x_j' &= x_0 - 2x_j, \\ x_l' &= x_l \qquad (l \neq 0, l \neq j), \end{aligned}$$

where $j = 1$ for I_{p_1} and $j = k + 1$ for $I_{p_{k+1}}$. Since the linear transformation i_j leaves L unaltered, (1) is established.

With the help of i_j we list those types of transformations in G which can be expressed as products of not more than five involutions I_P. The numbers given are the non-zero values of

$$x_0; x_1, x_2, \cdots :$$

(3)

	1; 0, 0, $\cdots$	(5.1.1)	14; 652
(1)	2; 1	(5.1.2)	20; 10, 522
(2)	4; 21	(5.2.1.1)	17; 844
(3.1)	5; 22	(5.2.1.2)	20; 10, 441
(3.2)	8; 421	(5.2.2.1)	14; 652
(4.1)	10; 522	(5.2.2.2)	20; 964
(4.2.1)	10; 441	(5.2.2.3)	26; 13, 642
(4.2.2)	13; 642	(5.2.3.1)	20; 8821
(4.2.3)	16; 8421	(5.2.3.2)	26; 12, 841
		(5.2.3.3)	29; 14, 842
		(5.2.3.4)	32; 16, 8421.

The numbers given in parentheses represent the genesis of the type opposite from earlier types.

4. The linear group g of types in the Cremona group G. We have seen in **3** that the characteristic $x_0, x_1, \cdots$ of the homaloidal web **3**(1) of the generic element of G may be obtained from the characteristic 1, 0, 0, $\cdots$ of the web of planes by a sequence with repetitions of the involutions $i_1, i_2, \cdots$ defined in **3**(2). We consider now the linear group g generated by these involutions i_j.

A particular generator i_1 of g has the period 2 and the determinant -1. So far as x_0, x_1 alone are concerned, i_1 has the invariant linear forms $x_0' - x_1' = x_0 - x_1$ and $x_0' - 3x_1' = -(x_0 - 3x_1)$. By combining the squares of these to eliminate x_0x_1 we find the invariant quadratic form $x_0^2 - 3x_1^2$. Hence

(1) *The linear group g generated by involutions i_j, which permutes the characteristics x of the Cremona elements in G, has the invariant linear and quadratic forms*

$$L = x_0 - x_1 - x_2 - \cdots,$$

$$Q = x_0^2 - 3x_1^2 - 3x_2^2 - \cdots.$$

For characteristics x of Cremona elements, $L = Q = 1$.

By solving the equations $L = Q = 1$ for given x_0 we can again find the webs given in the table of **3**(3) in the following order:

$$
(2) \qquad \begin{array}{ll}
1; 0\ 0 \cdots & 8; 4\ 2\ 1 \cdots \\
2; 1\ 0 \cdots & 10; 5\ 2\ 2 \cdots \\
4; 2\ 1 \cdots & \quad 4\ 4\ 1 \cdots \\
5; 2\ 2 \cdots & 11; 6\ 1\ 1\ 1\ 1\ 0 \cdots \\
7; \ldots\ldots & \ldots\ldots\ldots\ldots .
\end{array}
$$

However not all of the positive integer solutions x of the equations $L = Q = 1$ yield geometrically existent Cremona webs. The last case in (2) is an example. Let $x_1, x_2, x_3, \cdots$ be so arranged that $x_1 \geqq x_2 \geqq x_3 \geqq \cdots$. Then $3x_1x_2 \geqq 3x_2^2$, $3x_1x_3 \geqq 3x_3^2, \cdots$. Hence the difference $3x_1(L - 1) - (Q - 1) = 0$ yields the inequality $3x_1(x_0 - 1) \geqq x_0^2 - 1$. Thus, if $x_0 > 1, 3x_1 \geqq x_0 + 1$ and $3x_1 > x_0$. Now i_1 applied to the characteristic x yields $x_0' = 2x_0 - 3x_1 < x_0$. Thus the value x_0 of the characteristic x can always be reduced and eventually reduced to the $x_0 = 1$ of the web of planes unless in the process negative values of some of the $x_1, x_2, \cdots$ appear. We have already imposed the restriction that the values of x_i be positive or zero, i.e., that $x_i \geqq 0$ $(i = 1, 2, \cdots)$. These inequalities under the involutions i_j become successively $x_0 - 2x_i \geqq 0$, $2x_0 - 3x_i - 2x_j \geqq 0$, etc., the number of inequalities becoming infinite if as many as three variables x_1, x_2, x_3 appear. Hence

(3) *Every solution in positive integers of the equations $L = Q = 1$ represents a geometrically existent Cremona web unless the solution fails to satisfy some one inequality in the set which is conjugate to the initial inequalities $x_i \geqq 0$ under the group g. The group g on x_0, x_1, x_2, x_3 alone is infinite and the set of conjugate inequalities is likewise infinite.*

Let g_ρ be the subgroup of g which is generated by $i_1, i_2, \cdots, i_\rho$ alone. To complete the proof of (3) we have yet to show that g_3 is infinite and also that under g_3 the number of conjugate inequalities mentioned is also infinite. We observe that the involution i_1 has a single invariant linear space, $x_0 - 3x_1$, with a multiplier -1. Thus i_1 is a harmonic perspectivity determined by the point $1, 1, 0, 0 \cdots$ and its polar space as to Q. It is thus convenient to represent i_1 by this polar space and we prove that

(4) *In the g_3 generated by i_1, i_2, i_3 these three involutions are in an infinite conjugate set of generating involutions represented by the linear spaces*
$A_l\,(k\,\pm) \equiv (3k \pm 1)x_0 - 3(k \pm 1)x_l - 3kx_m - 3kx_n (l, m, n = 1, 2, 3; k \geqq 0)$, *and the group g_ρ $(\rho \geqq 3)$ is infinite.*

For, $x_0 - 3x_1$ is transformed by i_2 into $2x_0 - 3x_1 - 3x_2$ and this by i_1 into $x_0 - 3x_2$, whence i_1 and i_2 are conjugate even in g_2. These representative linear spaces are all of the generic form given in (4). We find that i_1 interchanges $A_1\,(k\pm)$, and also interchanges $A_2(k\pm)$ with $A_3([k \pm 1]\mp)$. Beginning then with any form of the set such as $A_1(k+)$ we apply i_1 to get $A_1(k-)$. Then i_2 applied to $A_1(k\pm)$ yields $A_3([k - 1]+)$ and $A_3([k + 1]-)$, and i_3 applied to these

yields $A_3([k-1]-)$ and $A_3([k+1]+)$. Also i_1 applied to $A_3([k-1]+)$ yields $A_2(k-)$, and i_2 applied to this yields $A_2(k+)$. Thus from $A_1(k+)$ we can get $A_l(k\pm)$, $A_l([k-1]\pm)$, and $A_l([k+1]\pm)$. This proves the theorem (4).

We examine finally the left member of the inequalities in (3). The form x_1 is unaltered by i_2 and i_3, but i_1 carries it into $x_1 + A_1(0+)$. In fact we can write i_1 as $x_0' = x_0 + A_1(0+)$, $x_1' = x_1 + A_1(0+)$. It is clear then from (4) that all the conjugates of x_1 can be expressed as x_1 plus a sum of forms A_l, the number of forms A_l being the number of times i_1 was used in forming the conjugate. We consider the element $e = i_1i_2i_1i_3$. Since $i_1i_2i_1$ is the transform of i_2 by i_1, it, along with i_3, is a member of the conjugate set of generating involutions. The product e is cyclic and of infinite period. In fact e transforms x_1 into $x_1 + A_1(1-) + A_2(1-)$. Moreover e transforms $A_1(k-) + A_2(k-)$ into $A_1([k+1]-) + A_2([k+1]-)$. Hence e^j transforms x_1 into $x_1 + \sum_{i=1}^{i=j} \{A_1(i-) + A_2(i-)\}$. These, for all values of j, are distinct, whence e has an infinite period. Hence

(5) *Under g_3 each of x_1, x_2, x_3 gives rise to an infinite set of conjugates.*

A question which frequently yields interesting results is that of the *symmetric* types in G—the types for which all the positive x_i have the same value. These are given by the equations $x_0 - \rho x_1 = 1$, $x_0^2 - 3\rho x_1^2 = 1$. For $\rho = 0$ we have the case $x_0 = 1$, the identical collineation. Otherwise $x_1(3-\rho) = 2$ and $x_1 = 1$, or $x_1 = 2$. Thus we find only the two cases: 2; 1 and 5; 2 2 given in the tables above.

5. Relation between the generic elements of g and of G. The generic element of g generated by $i_1, i_2, \cdots, i_\rho$ has the form

$$(1) \qquad g_1: \quad \begin{aligned} x_i' &= \alpha_{i0}x_0 - \alpha_{i1}x_1 - \cdots - \alpha_{i\rho}x_\rho \qquad (i = 0, 1, \cdots, \rho), \\ x_j' &= x_j \qquad (j > \rho), \end{aligned}$$

where the α's are positive integers or zero. Due to the invariance of Q and L the coefficients α satisfy the relations:

$$(2) \qquad \begin{aligned} &\sum_{i=1}^{i=\rho} \alpha_{ik} = \alpha_{0k} - 1 \qquad (k = 0, 1, \cdots, \rho); \\ &3\sum_{i=1}^{i=\rho} \alpha_{i0}^2 = \alpha_{00}^2 - 1; \quad 3\sum_{i=1}^{i=\rho} \alpha_{ik}^2 = \alpha_{0k}^2 + 3 \qquad (k = 1, \cdots, \rho); \\ &3\sum_{i=1}^{i=\rho} \alpha_{ik}\alpha_{il} = \alpha_{0k}\alpha_{0l} \qquad (k, l = 0, 1, \cdots, \rho; k \neq l). \end{aligned}$$

These conditions are sufficient to verify that the inverse of g_1 is

$$(3) \qquad g_1^{-1}: \quad \begin{aligned} x_0' &= \alpha_{00}x_0 - 3\alpha_{10}x_1 - \cdots - 3\alpha_{\rho 0}x_\rho, \\ x_i' &= (\alpha_{0i}/3)x_0 - \alpha_{1i}x_1 - \cdots - \alpha_{\rho i}x_\rho \qquad (i = 1, \cdots, \rho), \\ x_j' &= x_j \qquad (j > \rho), \end{aligned}$$

since $g_1 g_1^{-1}$ is the identity by virtue of the relations (2). This g_1^{-1} also leaves Q, L unaltered and thus we obtain a further set, complementary to (2), of relations on the coefficients α:

$$\sum_{j=1}^{j=\rho} \alpha_{0j} = 3\,\alpha_{00} - 1\,, \qquad \sum_{j=1}^{j=\rho} \alpha_{ij} = 3\,\alpha_{i0} - 1 \qquad (i = 1, \cdots, \rho)\,;$$

$$(4) \qquad \sum_{j=1}^{j=\rho} \alpha_{0j}^2 = 3\,(\alpha_{00}^2 - 1)\,, \qquad \sum_{j=1}^{j=\rho} \alpha_{ij}^2 = 3\,\alpha_{i0}^2 - 1 \qquad (i = 1, \cdots, \rho)\,;$$

$$\sum_{j=1}^{j=\rho} \alpha_{ij}\,\alpha_{kj} = 3\,\alpha_{i0}\,\alpha_{k0} \qquad (i, k = 1, \cdots, \rho;\, i \neq k)\,.$$

We now prove the theorem:

(5) *To the generic element g_1 of g there corresponds a Cremona transformation G_1 in G with C as an F-curve of the first kind for both the direct and inverse transformations, with $p_1, \cdots, p_\rho$ as isolated F-points, and with $q_1, \cdots, q_\rho$ as isolated F-points of the inverse. This element G_1 transforms the web of planes into the homaloidal web $(C^{\alpha_{00}-1} q_1^{3\alpha_{10}} \cdots q_\rho^{3\alpha_{\rho 0}})^{3\alpha_{00}-2}$. The P-surface of C is*

$$(C^{2\alpha_{00}-3} q_1^{6\alpha_{10}} \cdots q_\rho^{6\alpha_{\rho 0}})^{6(\alpha_{00}-1)}\,,$$

and the P-surface of p_j is $(C^{\alpha_{0j}/3} q_1^{\alpha_{1j}} \cdots q_\rho^{\alpha_{\rho j}})^{\alpha_{0j}}$.

The theorem being true for $I_1, \cdots, I_\rho$, we have only to show that it remains true for products. Let I_1 be the generating involution with F-point at q_1 which sends $q_2, \cdots, q_\rho$ into $q_2', \cdots, q_\rho'$. For convenience let $q_1 = q_1'$. Then the product $G_1 I_1$ transforms the web of planes into the transform of the homaloidal web described in (5) by I_1. This transform has the order $4(3\alpha_{00} - 2) - 6(\alpha_{00} - 1) - 9\alpha_{10} = 3(2\alpha_{00} - 3\alpha_{10}) - 2 = 3\alpha_{00}' - 2$. It has, for multiplicity on C, the value $3\alpha_{00} - 2 - (\alpha_{00} - 1) - 3\alpha_{10} = (2\alpha_{00} - 3\alpha_{10}) - 1 = \alpha_{00}' - 1$. It has, for multiplicity at q_1', the value $3(3\alpha_{00} - 2) - 6(\alpha_{00} - 1) - 6\alpha_{10} = 3(\alpha_{00} - 2\alpha_{10}) = 3\alpha_{10}'$. It has, for multiplicity at q_2', the value $3\alpha_{20} = 3\alpha_{20}'$. Thus the web $G_1 I_1$ has the form given in (5) for the values α'. But the values α' defined above are precisely the α''s of the product $g_1 i_1$ [cf. (1) and **3**(2)]. To complete the proof a similar check must be made for the P-surface of C by transforming the surface given in (5) by I_1; for the P-surface of p_1; and for the P-surface of p_2. It turns out as above that these transforms can be read off as in (5) from the coefficients α' of the product $g_1 i_1$. It is necessary also to carry out the same argument for the product $G_1 I_{\rho+1}$, $I_{\rho+1}$ being the generating involution with isolated F-point at $q_{\rho+1}$. We obtain the same check with the product $g_1 i_{\rho+1}$ and theorem (5) is established.

6. The transformation of particular surfaces or of linear systems of surfaces by elements of G. We define the *characteristic* of a surface or linear system, S, with respect to the element G_1 of G described in (5) to be a set of numbers

$$y,\, y_0,\, y_1,\, y_2,\, \cdots,\, y_\rho$$

of which y is the order of S, y_0 the multiplicity of S on C, and $y_1, y_2, \cdots$ the multiplicity of S at $p_1, p_2, \cdots$. Then S is transformed by G_1 into a system S' with characteristic $y', y'_0, y'_1, y'_2, \cdots$ with respect to the F-system $C, q_1, q_2, \cdots$ of G_1^{-1}. From the properties of G_1 given in **5**(5) we find at once that

$$\begin{aligned}
y' &= (3\alpha_{00} - 2)y - 6(\alpha_{00} - 1)y_0 - \alpha_{01}y_1 - \cdots - \alpha_{0\rho}y_\rho , \\
y'_0 &= (\alpha_{00} - 1)y - (2\alpha_{00} - 3)y_0 - (\alpha_{01}/3)y'_1 - \cdots - (\alpha_{0\rho}/3)y_\rho , \\
y'_i &= 3\alpha_{i0}y - 6\alpha_{i0}y_0 - \alpha_{i1}y_1 - \cdots - \alpha_{i\rho}y_\rho \qquad (i = 1, \cdots, \rho) , \\
y'_l &= y_l \qquad (l > \rho) .
\end{aligned}$$

We observe first that $y' - 3y'_0 = y - 3y_0$ and secondly that

$$\begin{aligned}
(y' - 2y'_0) &= \alpha_{00}(y - 2y_0) - \alpha_{01}(y_1/3) - \cdots - \alpha_{0\rho}(y_\rho/3) , \\
y'_i/3 &= \alpha_{i0}(y - 2y_0) - \alpha_{i1}(y_1/3) - \cdots - \alpha_{i\rho}(y_\rho/3) \qquad (i = 1, \cdots, \rho) , \\
y'_l/3 &= y_l/3 \qquad (l > \rho) .
\end{aligned}$$

From this there follows that

(1) *If in the linear transformation* **5**(1) *we make the change of variable* $x_0 = y - 2y_0$, $3x_i = y_i$ $(i = 1, 2, \cdots)$, *then the new transformation, coupled with* $y' - 3y'_0 = y - 3y_0$, *yields the linear transformation on the characteristic* y, y_0, y_i *which is effected by the Cremona transformation* G_1 *of* **5**(5).

It is clear that the infinite Cremona group G here defined has an arithmetic theory which is precisely parallel to that of the infinite ternary Cremona group.

UNIVERSITY OF ILLINOIS.

COLLECTIONS FILLING A PLANE

By J. H. Roberts

Introduction. In 1928 the author showed[1] that there exists an upper semi-continuous collection G filling a plane S such that every element of G is a bounded continuum not separating S. Later he stated[2] that the elements of G could all be taken to be bounded continuous curves. If M is a bounded continuous curve lying in a plane S and not separating S, either M is an arc or M contains a triod.[3] But any collection of mutually exclusive triods lying in a plane[3] is necessarily countable. Consequently, all of the elements of the collection G of continuous curves filling S, except possibly a countable number, are arcs. In view of this result it seemed likely that there existed an upper semi-continuous collection G filling S such that *every* element of G was an arc. In fact, the author has since stated[4] erroneously that such is the case. The principal object of the present paper is to prove that *there does not exist an upper semi-continuous collection G of arcs filling a plane S.* In view of this result, the fact that there is a collection G, every element of which is a bounded continuous curve not separating S, becomes of more interest, and accordingly an example of such a collection G is given.

Definition. A collection G of closed point sets lying in a metric space is said to be *upper semi-continuous*[5] if for each element g of G and each positive e there exists a positive d such that if x is an element of G and $l(x, g) < d$, then $u(x, g) < e$.

Definition. The element g of G is a *limit element* of a subcollection K of G if for every positive e there is an element x of K distinct from g such that $u(x, g) < e$.

Received November 15, 1935.

[1] Fundamenta Mathematicae, vol. 14 (1929), pp. 96–102.

[2] This result was presented to the North Carolina Academy of Sciences, May, 1934, but no published statement of it has appeared.

[3] A triod is the sum of three arcs AP_1, AP_2 and AP_3, each pair having only A in common. Cf. R. L. Moore, *Foundations of Point Set Theory*, Theorem 71, p. 250, and Theorem 75, p. 254. Theorem 75 is stated for a closed and compact set, but the present result obviously follows, since the plane is the sum of a countable number of such sets.

[4] See abstract #196, Bull. Amer. Math. Soc., vol. 41 (1935), p. 330.

[5] R. L. Moore, *Concerning upper semi-continuous collections of continua*, Trans. Amer. Math. Soc., vol. 27 (1925), pp. 416–428. If M is a point set and P is a point, then by $l(P, M)$ is meant the lower bound of the distances from P to all the different points of M. If M and N are point sets, then by $l(M, N)$ is meant the lower bound of the values $l(P, N)$ for all points P of M, while by $u(M, N)$ is meant the upper bound of these values for all points P of M. It is to be observed that $u(M, N)$ may be different from $u(N, M)$, while $l(M, N) = l(N, M)$. The quantities $l(M, N)$ and $u(M, N)$ are called the lower, respectively upper, distances of M from N.

DEFINITION. A collection G of point sets is said to *fill* a space S if every element of G is a subset of S and every point of S belongs to some element of G.

1. Let G denote an upper semi-continuous collection filling some metric space S. Let n be a positive integer and suppose d_n is a domain in S having the following properties: (1) $d_n = R_1 + R_2 + \cdots + R_k$, where R_i is a domain and[6] $\delta(R_i) < 1/n$, (2) d_n covers some element of G, but (3) $R_1 + \cdots + R_{i-1} + R_{i+1} + \cdots + R_k$ $(i = 1, 2, \cdots, k)$ does not cover any element of G. Let e_n be the set of elements of G covered by d_n. Then e_n is[7] a domain in G. Let E_n be the sum of all such domains e_n. Clearly $E_n \supset E_{n+1}$. Let G_1 be the inner limiting set (= G_δ set) common to $E_1, E_2, E_3, \cdots$.

THEOREM 1. *If the elements of G are closed and compact, then G_1 is a continuous*[8] *collection. Furthermore, if S is complete, then G_1 is maximal with respect to the property of being a continuous subcollection of G, and is dense in G.*

Suppose G_1 is not a continuous collection. There exist elements $g, g_1, g_2, g_3, \cdots$ such that $\lim_{n\to\infty} l(g_n, g) = 0$ but $u(g, g_n)$ does not approach zero as $n \to \infty$. There exists a positive e and a sequence $g'_1, g'_2, g'_3, \cdots$ such that $u(g, g'_n) > e$ and $g'_n = g_m$ for some m. For each n there is a point P_n of g such that $l(P_n, g'_n) > e$. There exists a limit point P of $P_1 + P_2 + P_3 + \cdots$, and there is an infinite subsequence $h_1, h_2, h_3, \cdots$ of $g'_1, g'_2, g'_3, \cdots$ such that $l(P, h_n) > e/2$ for every n.

Now choose a positive integer m such that $1/m < e/2$. Since g is in G_1, there is some domain d_m covering g and having the properties (1) and (3) as well. For some j the element h_j is covered by d_m. Set $d_m = R_1 + R_2 + \cdots + R_k$, these being domains of diameter $< 1/m < e/2$. For some integer r $(r \leqq k)$ the domain R_r contains P. Then R_r contains no point of h_j; i.e., h_j is covered by $R_1 + \cdots + R_{r-1} + R_{r+1} + \cdots + R_k$. This contradicts property (3) of d_m. Hence it has been proved that G_1 is a continuous collection.

Suppose now that S is complete. It is to be shown that G_1 is dense in G. Let D_1 be any domain of elements of G, and let g_1 be a particular element in D_1. One can take a finite set of domains in S whose sum covers g_1 and throw out domains of this set until what is left covers some element of G but no more domains can be omitted and that property retained. Thus there exists a domain d_1 having properties (1), (2) and (3) such that[9] $\bar{d}_1 \subset D_1^*$ and $u(d_1, g_1) < 1$. Let D_2 be the set of elements covered by d_1 and let g_2 be a particular such element.

[6] If X and Y are points of a metric space, then $\delta(X, Y)$ will denote their distance apart. More generally, if R is a point set, then $\delta(R)$ will denote the diameter of R.

[7] Moore, loc. cit., Theorem 1.

[8] The collection K is said to be *continuous* provided that for each element g and sequence $g_1, g_2, g_3, \cdots$ of elements of K such that $l(g_n, g) \to 0$ as $n \to \infty$, it follows that both $u(g_n, g)$ and $u(g, g_n) \to 0$ as $n \to \infty$. Thus in addition to being upper semi-continuous, a convergent sequence of elements of K has a whole element of K as its limiting set.

[9] If D is a set of elements of G, then D^* will denote the point set in S obtained by adding together all elements of D.

As above, there exists a domain d_2 having properties (1), (2) and (3), such that $d_2 \subset D_2^*$ and $u(d_2, g_2) < 1/2$. This process can be continued indefinitely. Thus there exist elements $g_1, g_2, g_3, \cdots$ of G and domains $d_1, d_2, d_3, \cdots$ in S such that (a) the domain d_n has properties (1), (2) and (3), and in addition covers g_n, (b) $\bar{d}_{n+1} \subset d_n$, and (c) $u(g_{n+i}, g_n) < 2/2^n$. It remains to show that there is an element g of G which is a subset of every d_n. For this it will be sufficient to show that there is some element g which is either g_n for infinitely many values of n or is a limit element of the set of elements $g_1 + g_2 + g_3 + \cdots$.

Clearly there exists an infinite sequence of domains in S, $E_1, E_2, E_3, \cdots$ such that (a) $E_n \subset d_n$, (b) $\delta(E_n) < 2/2^n$, (c) E_n contains a point P_n of g_n and (d) E_n and E_{n+1} have a point in common. The sequence $P_1, P_2, P_3, \cdots$ satisfies the Cauchy convergence condition, and since S is complete, there is a point P which is the sequential limit point of this sequence. Let g be the element of G which contains P. Then since $g_m (m \geqq n)$ is contained in the domain d_n, it follows that g is in $\bar{d}_n$, for every n. Hence g is covered by $d_1, d_2, d_3, \cdots$ and is, therefore, an element of G_1 lying in D_1.

It remains to show that G_1 is maximal with respect to the property of being a continuous collection. Let g be an element of G and suppose $G_1 + g$ is a continuous collection. Let n be given. There exists a domain c_n covering g such that (a) $c_n = K_1 + K_2 + \cdots + K_k$, where K_i is a domain in S of diameter less than $1/n$ and (b) for each i there is a point P_i in g and in K_i such that P_i is neither in nor on the boundary of K_j if $j \neq i$. There exists a positive ϵ such that every distance $l(P_i, \bar{K}_1 + \cdots + \bar{K}_{i-1} + \bar{K}_{i+1} + \cdots + \bar{K}_k)$ is greater than ϵ. Since $G_1 + g$ is continuous, there exists a δ_ϵ such that if h is in G_1 and $l(h, g) < \delta_\epsilon$, then $u(g, h) < \epsilon$. Then h has a point in K_i for every i $(i \leqq k)$. Let R_i be the subset containing every point of K_i at a lower distance less than δ_ϵ from g, and set $d_n = R_1 + R_2 + \cdots + R_n$. Then d_n covers g, and every R_i is of diameter less than $1/n$. Furthermore, if some R_i is omitted, the remaining domain covers no element of G_1; and since G_1 is dense in G, it can cover no element of G. Thus d_n has properties (1), (2) and (3). Therefore g is in G_1. This completes the proof of Theorem 1.

2. Definitions. Suppose $g_1, g_2, g_3, \cdots$ is a sequence of arcs converging to an arc g. Let AB be any subarc of g. Suppose that for every positive ϵ there exists an m such than if $n > m$, then g_n contains two mutually exclusive subarcs, each containing points at a distance less than ϵ from A and points at a distance less than ϵ from B. Then the subarc AB of g will be said to be *approached doubly* by the sequence $g_1, g_2, g_3, \cdots$. If no subarc of g is approached doubly by the sequence $g_1, g_2, g_3, \cdots$ which converges to g, then g is said to be *approached equi-continuously*[10] by the sequence $g_1, g_2, g_3, \cdots$.

Theorem 2. *If S is complete and G is an upper semi-continuous collection of*

[10] R. L. Moore, *Concerning certain equicontinuous systems of curves*, Trans. Amer. Math. Soc., vol. 22 (1921), definition, p. 42.

arcs filling S, there is a subcollection G_2 of G_1 (G_1 as defined in §1) *such that* (1) *G_2 is dense in G_1 (and therefore dense in G) and* (2) *every element g of G_2 is approached equi-continuously by every sequence of elements of G_1 converging to g.*

The author's detailed proof of Theorem 2 is long, and in some respects similar to the proof of Theorem 1. Therefore only an outline of the proof will be given. The following lemma is first established.

LEMMA 2.1. *Let ϵ be a positive number and let D be a domain in G. Let g be an arc of G_1 in D which contains n, but not $n + 1$, disjoint arcs each of diameter greater than ϵ; and suppose that g contains a subarc of diameter greater than ϵ which is approached doubly by some sequence of elements of G_1. Then there is an element g' of G_1 in D which contains $n + 1$ disjoint subarcs each of diameter greater than ϵ.*

With the help of the above lemma and processes described in the proof of Theorem 1, we next arrive at

LEMMA 2.2. *If ϵ is a positive number and D is a domain in G, there is a domain E in G such that $\bar{E} \subset D$, and if g is any element of G_1 in E, then no subarc of g of diameter greater than ϵ is approached doubly by any sequence $g_1, g_2, g_3, \cdots$ of elements of G_1 converging to g.*

Theorem 2 is then proved by showing the existence of a sequence of domains $E_1, E_2, E_3, \cdots$ lying in an arbitrary domain D of G such that (a) $\bar{E}_{n+1} \subset E_n$, (b) there is an element of G_1 common to $E_1, E_2, E_3, \cdots$, and (c) the domain E_n has the property of the domain E of Lemma 2.2 with ϵ equal to $1/n$.

3. The next three sections will be devoted to the proof of the following

THEOREM 3. *There does not exist an upper semi-continuous collection of arcs filling a plane.*

The proof is indirect. A contradiction is reached at the end of §5.

Suppose, then, that G is an upper semi-continuous collection of arcs filling a cartesian plane S. Let a and b denote mutually exclusive closed subsets of S. Let n and k be positive integers. Suppose D is a simple chain of domains such that (1) every domain of the chain D is of diameter $< 1/k$, (2) the sum of the domains of the chain D covers some element of G, and (3) there do not exist n subchains of D having at most end domains in common and each containing both a point of a and a point of b. Let $H_{nk}(a, b)$ be the set of all elements of G, each of which is covered by some such chain D. Let $K_{nk}(a, b)$ be $G - H_{nk}(a, b)$. Let $H_n(a, b)$ and $K_n(a, b)$ be defined as follows:

$$H_n(a, b) = H_{n1} \cdot H_{n2} \cdot H_{n3} \cdot \cdots ,$$

$$K_n(a, b) = K_{n1} + K_{n2} + K_{n3} + \cdots .$$

Then, as H_{nk} is a domain, it follows that K_{nk} is closed. Thus H_n is a G_δ set and K_n is an F_σ. It might be noticed that H_n contains just those elements of G which do not contain as many as n mutually exclusive segments each with end points on a and b, respectively.

4. We now specialize the sets a and b. Let $\rho(x, y)$ denote a distance function defined over G with respect to which G is a cartesian plane.[11] Let g be an element of G_2. Let J be a simple closed curve enclosing g. Let a_1, a, and b denote mutually exclusive arcs crossing g, having end points on J but otherwise lying within J, each having only one point on g and such that a separates a_1 and b in J plus its interior. There exists an ϵ such that (1) if h is in G and $\rho(h, g) \leqq \epsilon$, then h lies within J and cuts both a_1 and b, but (2) if h is in G_1 and $\rho(h, g) \leqq \epsilon$, then h does not cut a_1, or b, on both sides of g, and does not contain two mutually exclusive segments each having end points on a_1 and b, respectively. Let N denote the subcollection of G consisting of all elements h such that $\rho(h, g) \leqq \epsilon$. The following can now be established.

LEMMA 4.1. *The point set*[9] $[N \cdot K_2(a, b)]^*$ *contains a domain.*

Of the two domains into which a_1 divides the interior of J, let E denote the one such that $\bar{E} \supset a + b$. Let M be the collection of all maximal connected subsets of $h \cdot \bar{E}$ for all elements h of N. The collection M is upper semi-continuous and each of its elements is an arc or a single point. Each element of M contains a point of a_1. Obviously M^* contains a connected domain D which contains no point of g but does have on its boundary points of each of the segments AC and CB, where ACB is the arc a_1, C being on g. Then of the elements of M having points in D, some cut the segment AC, some BC, and all cut AC or BC. Since M is upper semi-continuous (and D connected), there is some element h_1 of M with a point in D such that h_1 cuts each of the segments AC and CB. Let g_1 be that element of M which is a subset of g and cuts the arc b. There is an element k of M which is a subset of an element of G_1 and which cuts[12] AC in a point T, and is such that k is not separated from g_1 in $\bar{E}$ by h_1. Then there is an element h_2 of M which cuts both AC and CB, and does separate k from g_1 in $\bar{E}$. Let E_1 be the set of all points P of $\bar{E}$ such that P can be joined in $\bar{E}$ to h_1 by an arc not intersecting h_2 and P can be joined to h_2 in $\bar{E}$ by an arc not intersecting h_1. Then E_1 is connected, has boundary points on AC and on CB. Hence some element h_3 of M having points in E_1 cuts both of the segments AC and CB. Let M_1 be the set of all elements of M which cut both of the segments AC and CB and contain points of $\bar{E}_1$. The elements of M_1 are obviously in a linear order, as one of every pair separates the other in $\bar{E}$ from g_1. We may say h_1 precedes all other elements of M_1 and h_2 follows all other elements of M_1. But M_1 is a closed collection. Furthermore, between any two elements of M_1 (as, for example, between h_1 and h_2) there is a third element of M_1. From all this it follows that M_1 is an arc of elements from h_1 to h_2. There is a first element l of this arc which cuts the segment TC. Then l also cuts AT (and also, of course, CB). The arc l contains a subarc $pqrst$, where p, r, and t are on a_1, and q and s are in $\bar{E}$ on the non-a side of the arc b. There is a domain L with the following properties: (1) L lies in E between a_1 and a, (2) L is bounded by a simple closed curve which contains as a subset a subarc of l containing r in its interior, and (3) the diameter of L is so small that if P is a point of L and g_P

[11] R. L. Moore, footnote 5.

[12] There is such a k cutting AC or CB, and to be explicit, we may suppose it is AC.

is the element of G containing P, then g_P is in N, and therefore cuts a_1. Then L is the domain desired, being in $[N \cdot K_2(a, b)]^*$. For starting on the arc g_P from the point P of L and going in either direction, one intersects the arcs a, b and a again before it is possible to intersect a_1. Thus g_P is in $K_2(a, b)$. This completes the proof of 4.1.

5. Since $K_2^*(a, b)$ (a and b as defined in the previous section) contains the domain L, it follows[13] that for some i_2 the closed set $K_{2i_2}^*(a, b)$ contains a bounded domain D_2 which is a subset of L. It may be that no matter how small the positive number ϵ_3 is, there exist arcs a_3 and b_3 similar to a and b (as regards cutting g and J) with $u(a_3, a) < \epsilon_3$ and $u(b_3, b) < \epsilon_3$ such that $K_3^*(a_3, b_3)$ contains a domain which is a subset of D_2. In this case choose ϵ_3 to be $1/8$ of $l(a, b)$ and select an integer i_3 such that $K_{3i_3}^*(a_3, b_3)$ contains a domain D_3 such that $\bar{D}_3 \subset D_2$.

Suppose this process can be continued and that (1) for every n the set $K_{ni_n}^*(a_n, b_n)$ contains a bounded domain D_n, (2) $\bar{D}_{n+1} \subset D_n$, (3) $u(a_n, a)$ and $u(b_n, b)$ are less than $1/4$ of $l(a, b)$, and (4) the arcs a_n and b_n cross g, and have end points on J. Then there is a point P common to $D_1, D_2, D_3, \cdots$. Let g_P be the arc of G which contains P, and let a' and b' be arcs crossing g (similar to a and b and a_n and b_n) such that (1) $l(a', b') > 1/8$ of $l(a, b)$, and (2) a', as well as b', separates every a_n from every b_n within J. Then every subarc of g_P which cuts a_n and b_n cuts a' and b'. Now g_P belongs to $K_n(a_n, b_n)$, and hence to $K_n(a', b')$ for every n. But then g_P contains infinitely many arcs spanning from a' to b', mutually exclusive closed sets. This is impossible, and we have arrived at the following:

Lemma 5.1. *There exist arcs a and b, an element g of G_2, a positive ϵ and integers n and k, and a simple closed curve J such that* (1) *J encloses g,* (2) *the arcs a and b are mutually exclusive, have their end points on J, otherwise lie within J, and cross g in points E and F, respectively, these being the only points of g on $a + b$,* (3) *every element of G which lies within J cuts both a and b, but no element h of G_1 which lies within J cuts either a or b on both sides of g, nor does h contain two mutually exclusive segments each having its end points respectively on a and b,* (4) *$K_{nk}^*(a, b)$ contains a domain D and every element of G with a point in $\bar{D}$ is within J, and* (5) *if a' and b' are arcs such that $u(a', a) < \epsilon$ and $u(b', b) < \epsilon$, then $K_{n+1}^*(a', b')$ does not contain a domain which has a point in common with D.*

Select definite arcs a' and b' distinct from a and b with properties similar to those obtaining for a and b and such that (1) in the interior of J the order is $aa'b'b$, and (2) $u(a', a) < \epsilon$ and $u(b', b) < \epsilon$. Let M be the collection of arcs of $K_{nk}(a, b)$ which contain points in D. Let L be the set of all elements of M which have exactly n distinct arcs spanning[14] a' to b', i.e.,

$$L = M \cdot [K_n(a', b') - K_{n+1}(a', b')].$$

[13] The theorem is as follows: If for every n the closed plane set M_n contains no domain, then $M_1 + M_2 + M_3 + \cdots$ contains no domain.

[14] A set Z of arcs will be called a set of *distinct arcs spanning a' to b'* if no two arcs of Z have more than one end point in common, and each arc of Z has its end points on a' and b', respectively.

It is merely a matter of notation to assume that D is the *maximal domain* which is a subset of M^*. The following assertion follows obviously from (5) of Lemma 5.1:

LEMMA 5.2. *The point set L^* is dense in D.*

Now let g_1 denote any element of L. Then g_1 contains at least n arcs spanning a to b, since g_1 is in $K_n(a, b)$. But since every arc from a to b contains a subarc from a' to b', it follows that, since g_1 is in L, it contains not more than n distinct arcs spanning a to b. Thus if A and B are the end points of g_1, there exist on g_1 points $A, Q_1, R_1, Q_2, R_2, \cdots, Q_n, R_n, B$ in the order written such that (1) the segment Q_iR_i contains no point of a or of b, (2) one of the points Q_i and R_i is on a, the other on b, and (3) any subarc of g_1 which contains both a point of a and a point of b contains the entire arc Q_iR_i, for some i.

There exists a simple chain N of connected domains irreducibly[15] covering g_1 such that (1) every domain of N is of diameter less than $1/k$, (2) for each i ($i = 1, 2, \cdots, n$) N contains a subchain C_i which irreducibly covers the arc Q_iR_i, (3) only one domain of the chain C_i cuts a and only one cuts b, (4) no domain of N contains points of both a and a', or of both b and b', and (5) neither of the two subchains of N consisting of all domains of N which contain a point of AQ_1, or R_nB, respectively, cuts both a' and b'.

LEMMA 5.3. *Every element of K_{nk} which is covered by N must contain a point in every region of C_i for every i ($i = 1, 2, \cdots, n$).*

The proof of Lemma 5.3 is immediate. For if h is an element of G which is covered by N but does not intersect every domain of C_i for some i, then h is covered by a subchain of N which does not contain as many as n subchains each spanning from a to b. Then h is in H_{nk}, and hence not in K_{nk}. (See §3 for definitions.)

There are now two cases to be disposed of.

Case 1. Suppose the point A (an end point of g_1) is in D.

We can let g_1 play the rôle of the arc g of §4 and apply the argument which proved Lemma 4.1, with scarcely a change, except in notation. The following lemma results:

LEMMA 5.4. *There exists an arc qrs which is a subarc of an element h of M such that*

(1) *q and s are on*[16] *a, and r is on b,*

(2) *if J denotes the simple closed curve qrs plus the subarc qs of a, then the interior of J and the segment qs of a are subsets of D,*

(3) *the interior of J contains points of a' and of b', but no point of a or of b, and*

(4) *J plus its interior is covered by the chain C_1 (the subchain of N covering the arc Q_1R_1).*

By Lemma 5.2 some element V of L has a point Z within J between b' and b. If a point P moves along V, starting at Z, then before it can cut the arc b it

[15] A set N of domains will be said to cover irreducibly a point set g_1 if N covers g_1, but no domain of the set N can be omitted and that property retained.

[16] The pairs (a, a') and (b, b') are essentially interchangeable, and so to avoid ambiguity it can be supposed that the notation has been properly assigned.

must cut b', a', a, a', and b'. But this is impossible for an arc V of L, which contains precisely n arcs spanning a to b and precisely n arcs spanning a' to b'.

Case 2. Neither end point of g_1 is in D. The subarcs AQ_1R_1 and Q_nR_nB of g_1 are in $S - D$, since every element of G (sufficiently close to g_1) cuts both a and b. It follows that i can be so chosen ($i < n$) that Q_iR_i contains no point of D, but $R_iQ_{i+1}R_{i+1}$ does contain a point of D. Due to symmetry with respect to a and b we can suppose that Q_i and R_{i+1} are on a, while R_i and Q_{i+1} are on b.

We again use an argument very similar to that given in §4, and prove that there exists an element h of G which is covered by N and contains a subarc $pqrst$ with the following properties: (1) some point P of the segment pq is in D, (2) p, r, and t are on a and are the only points of a on $pqrst$, (3) q and s are on b, and (4) p and r are in domains of the chain N containing R_{i+1} and Q_i, respectively. Now an element V of L can be chosen with $l(V, P)$ so small that V has points in every domain of C_i and C_{i+1}. Such an arc V cannot have points on both sides of $pqrst$ within C_i or C_{i+1} (since V is in L). Think of a point Z moving along V into the chain C_i (see Lemma 5.3). If V is taken close enough to h, and on the proper side, then Z will cut the arcs b, b', a', a', b', b, a' before it can cut the arc a, since it cannot cross $pqrst$. But then V has more distinct arcs spanning a' to b' than it has spanning a to b. This is impossible and the proof of Theorem 3 is complete.

6. In this section the following theorem is established.

THEOREM 4. *There exists an upper semi-continuous collection F filling a plane S such that every element of F is a bounded continuous curve not separating S.*

There will first be described a collection G of arcs such that G is an open curve and[9] G^* is a continuum M. The set M will be the common part of an infinite sequence of sets $M_1, M_2, M_3, \cdots$. Let M_1 be the set of all points with coördinates (x, y) such that $0 \leqq y \leqq 1$.

By a V, in the following description, will be meant an arc which is the sum of two intervals, with end points respectively on the lines $y = 0$ and $y = 1$, the intervals being subsets of lines whose slopes are in absolute value > 1. It will be said that the V opens upward, or downward, depending upon whether the end points of the V are on the line $y = 1$, or on the line $y = 0$. There exist two sets of V's, G_{11} and G_{12}, such that

(1) the elements of $G_{11} + G_{12}$ are mutually exclusive,

(2) the V's of G_{11} open upward, and those of G_{12} open downward,

(3) the number of V's in $G_{11} + G_{12}$ within any circle is finite, and

(4) as a point P moves along the line $y = 0$ it intersects a V of G_{11} between every two V's of G_{12}, and vice versa.

Let M_2 be M_1 minus all points of M_1 which lie *within* some V (i.e., points which lie in the pair of *smaller* angles which the two lines whose sum contains the V make). Next, there exist sets of V's, G_{21} and G_{22} such that if h and k are consecutive[17] V's of the set $G_{11} + G_{12}$ (whence one is in G_{11}—call it h—and the

[17] The linear order is obvious.

other in G_{12}), then between h and k there is precisely one element g_1 of G_{21} and one element g_2 of G_{22}, the order being h, g_2, g_1, k. Let M_3 be M_2 minus all points of M_2 within some V of $G_{21} + G_{22}$.

Clearly, then, there exists an infinite sequence of pairs of sets of V's $G_{11}, G_{12}; G_{21}, G_{22}; G_{31}, G_{32}; \cdots$ such that for every n the V's of the set K_n $[K_n = \sum_{i=1}^{n}(G_{i1} + G_{i2})]$ are mutually exclusive, are alternately (as one intersects them along the x-axis) open upward and downward, and such that every interval of the line $y = k$ $(0 \leqq k \leqq 1)$ neither intersecting nor lying within any V of the set K_n is of length less than $1/n$.

Let M_{n+1} denote M_n minus all points of M_n which lie within any V of $G_{n1} + G_{n2}$. Let M be the common part of $M_1, M_2, M_3, \cdots$. Let G be the collection $\sum_{i=1}^{\infty}(G_{i1} + G_{i2})$ plus all maximal connected subsets of $M - \sum_{i=1}^{\infty}(G_{i1} + G_{i2})$. Each of these maximal connected subsets is, clearly, an interval with end points on the lines $y = 0$ and $y = 1$, respectively. It is also seen that G is an open curve of arcs.

Now in my paper *Concerning atriodic continua*[18] I showed that if M is a continuum in a plane S which, for every positive number e, can be covered by a simple chain of connected domains all of diameter $< e$, there exists in S an uncountable set K of mutually exclusive continua each homeomorphic with M. This result can be extended to the case where M can be covered, for every positive e, by an *unbounded chain* of connected domains (i.e., a domain D_i for every integer i such that (1) D_i and D_{i+1} have a point in common but (2) D_i and D_{i+j} have no point in common if $j > 1$), all of diameter $< e$. Now clearly the continuum M defined above can be covered, for every positive e, by an unbounded chain of connected domains of diameter $< e$.

The following lemma will be stated without further proof.[19]

Lemma 6.1. *There exists a set T of topological transformations of the plane S into itself such that:*

1. *If A denotes the set of all triadic decimals k $(0 \leqq k \leqq 1)$ each of whose digits is 0 or 2, then for every k in A there is a transformation T_k of the set T, and every transformation of the set T is some T_k.*
2. *If we denote $T_k(M)$ by M_k and $T_k(G)$ by H_k (i.e., H_k is the upper semi-continuous collection of arcs of which M_k is the sum), then every M_k separates the plane into two domains, one containing every M_h such that $h < k$, the other every M_h such that $h > k$.*
3. *Every element of H_k is of diameter greater than one.*
4. *If k_1 and k_2 denote respectively the fractions (triadic) $n_1n_2 \cdots n_i0222 \cdots$*

[18] Monats. für Math. und Phys., vol. 37 (1930), pp. 223–230. The argument given proves more than is stated in the theorem.

[19] Part of the argument omitted is an easy generalization of that in my paper, ibid. The remainder is fairly obvious, graphically, but a detailed logical proof is tedious in the extreme.

and $n_1n_2 \cdots n_i2000 \cdots$, and D_i the domain of S complementary to $M_{k_1} + M_{k_2}$ which contains no M_k, then there exists a collection K_i of arc segments such that

(a) *K_i^* fills D_i,*

(b) *there is a one-to-one correspondence between the segments of K_i and pairs of corresponding elements of H_{k_1} and H_{k_2} (i.e., elements which are images under T_{k_1} and T_{k_2}, respectively, of the same element of G),*

(c) *if r is a segment of K_i and g_1 and g_2 are the corresponding elements of H_{k_1} and H_{k_2}, then the end points of r are end points of g_1 and g_2, respectively, unless g_1 and g_2 are V's (see definition of M), in which case the end points of r are the vertices of the V's, and the diameter of $g_1 + g_2 + r$ does not exceed the diameter of g_i $(i = 1, 2)$ by more than $1/i$, and finally,*

(d) *the collection L_i of continuous curves $g_1 + g_2 + r$ is upper semi-continuous and is an open curve of elements.*

Let k_1 be $.000 \cdots$ and k_2 be $.222 \cdots$. Let D denote the complementary domain of $M_{k_1} + M_{k_2}$ in S which is bounded by $M_{k_1} + M_{k_2}$. Let W be the collection consisting of (1) all elements of L_i $(i = 1, 2, 3, \cdots)$, and (2) all elements of H_k for every k of the set A which is either not periodic at all, or has a period greater than 1 (i.e., in k the digits 0 and 2 each occur infinitely many times). Now W fills D. Furthermore W is upper semi-continuous. For let h_1 and h_2 be distinct elements of W. There is a positive integer i such that $l(h_1, h_2) > 2/i$. Suppose the elements $g_1, g_2, g_3, \cdots$, all distinct from h, contain points $P_1, P_2, P_3, \cdots$ converging to a point P of h_1. Let g_n' be an element of some H_k which is a subset of g_n. Let Q be the set of all elements g_n such that $d(g_n) - d(g_n') > 1/i$, and let R be the set of all other elements g_n. Clearly h_2 contains no limit point of the sum of the elements of R. On the other hand, there exists an m such that every element of Q belongs to $L_1 + L_2 + \cdots + L_m$. Since L_j is upper semi-continuous for every j, the element h_2 contains no limit point of the sum of the elements of Q.

Now suppose Z is a topological transformation of the domain D into the plane S which does not decrease the distance between any pair of points of D. Let F denote the collection of images under Z of the elements of W. Then F is an upper semi-continuous collection filling the plane S such that (1) every element of S is either an arc, or the sum of three arcs a, b, and c, which make an H (i.e., a and b are mutually exclusive, and c has just its end points on $a + b$, these being interior points of a and b, respectively), and (2) every element of F is of diameter greater than 1.

DUKE UNIVERSITY.

ON CERTAIN ANALYTIC CONTINUATIONS AND ANALYTIC HOMEOMORPHISMS

By Arthur B. Brown

1. Introduction. We generalize to the case of n complex variables and one real variable a theorem of Severi[1] regarding analytic continuation, over a limited domain[2] in the $(2n + 1)$-space of the variables, of a function given analytic near the boundary B. The theorem states that if B is connected the continuation is possible. Severi proves the theorem only for the case that $n = 1$ and the domain is of simple type. We remove all restrictions as to simplicity of the domain and its boundary.

The similar theorem for a region in the $2n$-space of $n > 1$ complex variables is Osgood's[3] extension of a theorem of Hartogs.[4] Because of certain geometric difficulties which seem not to be fully met in Osgood's proof, we give a detailed proof of this theorem. The proof applies without essential modification to the case of meromorphic continuation.[5]

As an application, we prove in the case of n complex variables that if the connected boundary of a limited domain in the space undergoes an analytic homeomorphism with non-vanishing jacobian, the transformation can be continued analytically over the domain to yield an analytic homeomorphism of the domain and its boundary (Theorem 4.II). A somewhat similar result is obtained for the case of one real and n complex variables (Theorem 4.III).

2. Functions of n complex variables. The following is the Osgood form of the theorem of Hartogs.

Theorem 2.I. *Let $\mathfrak{R}$ be a limited domain with connected boundary B in the $2n$-*

Received October 16, 1935; presented to the American Mathematical Society, February 23, 1935.

[1] F. Severi, *Una proprietà fondamentale dei campi di olomorfismo di una variabile reale e di una variabile complessa*, Atti della Reale Accademia Nazionale dei Lincei, Rome, Rendiconti, (6), vol. 15 (1932), pp. 487–490. Our theorem is numbered 3.II.

[2] By a domain we mean an open set. A region is a connected open set. A limited point set is one of finite diameter.

[3] W. F. Osgood, *Lehrbuch der Funktionentheorie*, vol. 2, part I, Chapter 3, §11. We refer to the book as Osgood II.

[4] F. Hartogs, *Einige Folgerungen aus der Cauchyschen Integralformel bei Funktionen mehrerer Veränderlichen*, Sitzungsberichte der mathematisch-physikalischen Klasse der K. B. Akademie der Wissenschaften, München, vol. 36 (1906), pp. 223–241. Hartogs proves only that if a function is given defined over the entire region and boundary, analytic at the boundary and without removable singularities in the region, it is analytic in the region.

[5] Theorem 2.II. See Osgood II, Chapter 3, §13, and E. E. Levi, *Studii sui punti singolari essenziali delle funzioni analitiche di due o più variabili complesse*, Annali di Matematica, (3), vol. 17 (1910), pp. 61–87.

space of the n complex variables $x_1, \cdots, x_n$, $n > 1$, *and* $f(x_1, \cdots, x_n) \equiv f(x)$ *a function single-valued and analytic in a domain* $\mathfrak{S}$ *containing* B. *Then* f *admits a single-valued analytic continuation throughout* $\mathfrak{R} + \mathfrak{S}$.[6]

First we divide $2n$-space into $2n$-cubes of diameter less than the minimum distance from B to the boundary of $\mathfrak{S}$, determined by $(2n - 1)$-planes parallel to the coördinate planes. Let A be the set of all points on closed cubes which meet $R = \mathfrak{R} + B$. Since B is connected, R is connected; hence A is connected. If the points of A are removed from $2n$-space, there remains a set of one or more regions one and only one of which, say $\mathfrak{L}$, is not limited. The boundary of the latter is denoted by C, which must then be the locus of a $(2n - 1)$-cycle both orientable and (mod 2). The cells are faces, and edges of lower dimensions, of some of the cubes. We note that C is part or all of the boundary of A. Then C bounds a limited domain $\mathfrak{T} = 2n$-space minus $(\mathfrak{L} + C)$, so that $T_0 = \mathfrak{T} + C$ contains A, and $\mathfrak{T}$ contains R. Also, C is connected; otherwise it would be easy to show that A is not connected. Let $\mathfrak{E} = T_0 + \mathfrak{S}$. We shall prove that f can be continued analytically over all of $\mathfrak{E}$.

Let P be a point not in $\mathfrak{E}$. A $(2n - 1)$-sphere Σ with center at P will be said to be *reachable* if there exists a function $\phi(x)$, defined, single-valued and analytic over the part of $\mathfrak{E}$ outside Σ, with $\phi(x) = f(x)$ in the part of $\mathfrak{S}$ outside Σ and not in T_0. Now if it were impossible to continue over all of $\mathfrak{E}$, not all spheres with center at P would be reachable, and we could let Σ_0 be the $(2n - 1)$-sphere with center at P whose radius was the greatest lower bound of the radii of all the reachable spheres. By considering analytic continuations radially towards P, one proves easily that Σ_0 is itself reachable. Evidently no sphere smaller than Σ_0 can be reachable. Let us now suppose Σ_0 to exist, and show that a contradiction must arise.

Case I. *If* Q *is any point of* $\Sigma_0 \cdot C$ (*intersection of* Σ_0 *and* C), *then* $f(x) = \phi(x)$ *at the nearby points outside of* Σ_0. It then follows that $f(x)$ near Q gives a proper analytic continuation $\zeta(x)$ of $\phi(x)$ throughout a small sphere δ in $\mathfrak{S}$ with center at Q. Now suppose Q is on $\Sigma_0 \cdot \mathfrak{T}$. Then $\phi(x)$ is single-valued and analytic in the part outside Σ_0 of a neighborhood of Q. According to a theorem[7] resulting from the work of F. Hartogs and E. E. Levi,[8] $\phi(x)$ is analytically continuable, say by $\zeta(x)$ defined throughout a small spherical neighborhood δ of Q in $\mathfrak{T}$. We deduce easily that a sphere smaller than Σ_0 will be reachable, a contradiction. Hence if there is any such sphere Σ_0, Case I cannot hold.

Case II. *Not Case* I. There is at least one point Q_0 on C such that $f(x)$ and $\phi(x)$ are unequal at some points near Q_0 outside of Σ_0. Then none of the part of C near Q_0 can be outside of Σ_0, for if it were, $f(x)$ would equal $\phi(x)$ where both are defined near Q_0, as follows from the uniqueness of analytic continuation. Since

[6] $\mathfrak{R} + \mathfrak{S}$ is the set of points each of which is in at least one of the sets $\mathfrak{R}$ and $\mathfrak{S}$. Notations of topology will be as in S. Lefschetz, *Topology*, Amer. Math. Soc. Colloquium Publications, vol. 12, New York, 1930, (Lefschetz I).

[7] Osgood II, Chapter 3, §10, Zusatz.

[8] Loc. cit.

the cells of C as a complex are planar, it follows that none of them of positive dimension can meet Σ_0 near Q_0, so that Q_0 is an isolated point of $\Sigma_0 \cdot C$. The nearby points outside of or on Σ_0 must then belong to $\mathfrak{T}$, as otherwise we could not have Case II for Q_0.

In the following we use the property that certain closed loci are complexes in the sense of analysis situs, after a proper subdivision of the entire configuration into cells.[9]

Let F_0 denote the set obtained from $\Sigma_0 \cdot C$ by removing each of the points at which the distance from P on C has a maximum. Consider the class (S) of $(2n - 1)$-spheres with center at P such that S is in the class, if the maximal connected set containing Q_0 which is a subset of the part of C not interior to S also contains points of F_0. Let S_0 be the $(2n - 1)$-sphere with center at P whose radius is the least upper bound of the radii of the spheres of this class. Since C is closed, S_0 is itself in the class, and it is the largest sphere of the class.

Let D_1 be the closure of the maximal connected part of $C - C \cdot S_0$ which contains Q_0. Since C is a $(2n - 1)$-cycle (mod 2), the chain boundary E_1 of D_1 is on S_0, and hence is a $(2n - 2)$-cycle of S_0. Then D_1 plus each of the two $(2n - 1)$-chains of S_0 bounded by E_1 is a $(2n - 1)$-cycle, and hence bounds a limited chain in $2n$-space. Let H_1 denote the closed locus of that one of the two $2n$-chains which contains no cells interior to S_0. It is not hard to show that *no point of H_0 is outside Σ_0*.

Since the distance from P on C has a maximum at Q_0, the part of C near Q_0 is the part near Q_0 of the boundary of only one of the $2n$-cells used in determining A, and that one, say K, is in H_1, since no point of H_1 is outside Σ_0. Also the part not on K of a neighborhood of Q_0 must be in $\mathfrak{T}$, since the part outside Σ_0 of a neighborhood of Q_0 is in $\mathfrak{T}$. Consequently, if we let T_1 be the part not interior to S_0 of the set obtained by adding to T_0 those points of H_1 not already in T_0, *Q_0 is an interior point of T_1*. Let C_1 denote the closure of the part outside S_0 of the boundary of T_1.

We now consider the auxiliary problem of the analytic continuation over $T_1 + \mathfrak{S}$ of the values of $f(x)$ as given near C_1.

Let us again consider spheres Σ with center at P and radii larger than that of S_0. Proceeding as in the earlier part of our proof, suppose first that for the auxiliary problem Case II does not arise for any of these spheres. The analytic continuation can be carried as far as S_0, and whether or not Case II arises at S_0, we can have a single-valued analytic function $\psi(x)$ defined over a domain containing T_1. Since the part of C not inside S_0 is a complex, and S_0 is in class (S), we can obtain a curve on C joining Q_0 to a point Z of F_0 and not passing into the interior of S_0. Since there are points near Z and outside Σ_0 on C, hence on C_1, Z is on C_1. Therefore $\psi(x) = f(x)$ near Z. By the uniqueness of analytic

[9] B. O. Koopman and A. B. Brown, *On the covering of analytic loci by complexes*, Trans. Amer. Math. Soc., vol. 34 (1932), pp. 231–251 (Theorem 6.I and Lemma 3.I). For another treatment, see S. Lefschetz and J. H. C. Whitehead, *Analytical complexes*, ibid., vol. 35 (1933), pp. 510–517, and Lefschetz I.

continuation along the curve back to Q_0, $\psi(x)$ must equal $f(x)$ near Q_0. But $\psi(x)$ is analytic near Q_0 since Q_0 is an interior point of T_1, and it equals $\phi(x)$, where the latter is defined near Q_0. Consequently $f(x)$ and $\phi(x)$ must be equal where both are defined near Q_0, contrary to the hypothesis that Case II arises at Q_0.

It is thus seen that Case II must arise for the auxiliary problem for T_1, say at a point Q_1, with Q_1 outside of Σ_0. Let Σ_1 be the $(2n - 1)$-sphere with center at P and passing through Q_1, and $\phi_1(x)$ the single-valued analytic function defined over the part of $T_1 + \mathfrak{S}$ outside of Σ_1, with $\phi_1(x) = f(x)$ in the part of $\mathfrak{S}$ outside Σ_1 and not in T_1. Let F_1 be the set obtained from $\Sigma_1 \cdot C_1$ by removing each of the points at which the distance from P on C_1 has a maximum. We now distinguish between two further subcases.

Case IIa. *Point Q_1 cannot be connected to any point of F_1 by a curve on C_1 not passing interior to S_0.* We denote by D_2 the closure of the maximal connected part of $C_1 - C_1 \cdot S_0$ which contains Q_1. Next we determine a set H_2 bounded partly by D_2, just as H_1 was determined above, and add to T_1 those points of H_2 not already in T_1, calling the resulting set T_2. This will make Q_1 an interior point of T_2, and we then proceed as before in a new auxiliary problem, with the process of continuing analytically as far as S_0 if possible, thus obtaining a contradiction at Q_0.

Case IIb. *Point Q_1, while in Case II, is not in Case IIa.* We let S_1 be the $(2n - 1)$-sphere with center at P of maximum radius such that Q_1 can be joined to a point of F_1 by a curve on C_1 not passing into the interior of S_1. We proceed in this case as we did above in the first consideration of Case II, with Q_1, S_1, F_1, C_1 taking the rôles of Q_0, S_0, F_0 and C, respectively. At no later step of this treatment of Case II will it be necessary to consider any sphere smaller than S_1, which of course is at least as large as S_0.

The procedure is now clear. After a finite number of steps we must obtain a contradiction, because each Q_j is a 0-cell of the original C, and there is only a finite number of the latter. It follows that Case II cannot arise. Hence Theorem 2.I is true.

THEOREM 2.II. *Theorem* 2.I *remains true if the word "analytic" is replaced by "meromorphic".*[10]

Since the proof is exactly similar to that of Theorem 2.I, we omit it.

3. **Functions of one real and n complex variables.** We shall prove a theorem similar to Theorem 2.I. As a preliminary step we now extend that theorem to the case in which parameters are involved.

THEOREM 3.I. *Let $\mathfrak{R}$ be a limited domain, with connected boundary B, in the $2n$-space of the n complex variables $x_1, \cdots, x_n$, $n > 1$, and $f(x_1, \cdots, x_n, y_1, \cdots, y_p) \equiv f(x, y)$ a function single-valued and analytic in the cylindrical region for which (x) is in an open set $\mathfrak{S}$ containing B and (y) in a region $\mathfrak{V}$ in complex (y)-space. Then the analytic continuation of f over $\mathfrak{R}$ for each point (y) deter-*

[10] See footnote in introduction referring to this theorem.

mines a function analytic in $x_1, \cdots, y_p$ *in the cylindrical region for which* (x) *is in* $\mathfrak{R} + \mathfrak{S}$ *and* (y) *is in* $\mathfrak{V}$.

We begin as at the beginning of the proof of Theorem 2.I, and let C and $\mathfrak{T}$ be defined as in that proof. In the proof of Theorem 2.I we showed that for any fixed (y) in $\mathfrak{V}$, f has a unique analytic continuation $\phi(x, y)$ over all of $\mathfrak{E} = \mathfrak{T} + \mathfrak{S}$. Thus $\phi(x, y) = f(x, y)$ in $\mathfrak{S}$.

Let P with coördinates (x^0) be in $\mathfrak{T}$ and (y^0) be in $\mathfrak{V}$. We must prove that $\phi(x, y)$ is analytic at (x^0, y^0). Let ρ be the 2-plane in (x)-space parallel to the 2-dimensional x_1-plane and passing through P. The set $\rho \cdot \mathfrak{T}$ contains a finite number of regions on ρ, one of which, say $\mathfrak{U}$, contains P. Let $\mathfrak{W}$ be the region obtained by adding to $\mathfrak{U}$ all the isolated points of the boundary of $\mathfrak{U}$ on ρ. If L is the point set boundary of $\mathfrak{W}$ on ρ, then $\mathfrak{W} + L$ is easily shown to be a complex, and L is thus the locus of a 1-cycle. Of course L need not be connected.

Let $J(L)$ denote the cylindrical point set consisting of the points of (x)-space with $x_1 =$ any value on L, and $x_2, \cdots, x_n$ near $(x_2^0, \cdots, x_n^0)$. Since L is on C, $\phi(x, y)$ is analytic in a domain with (x) near $J(L)$ and (y) near (y^0). Let L_1 denote the projection of L on the x_1-plane. Now consider the function

$$\frac{1}{2\pi i}\int_{L_1} \frac{f(t, x_2, \cdots, x_n, y_1, \cdots, y_p)}{t - x_1}\, dt, \tag{3.1}$$

the integral being taken in the positive sense over L_1 as the boundary of the projection $\mathfrak{W}_1$ of $\mathfrak{W}$. Here $x_2, \cdots, x_n, y_1, \cdots, y_p$ are regarded as parameters, and the values taken are those of f with (x) near $J(L)$. Evidently (3.1) defines an analytic function of $(x_1, \cdots, x_n, y_1, \cdots, y_p)$ for (x, y) near (x^0, y^0). For fixed (y), ϕ is known to be analytic in (x) over $\mathfrak{E}$. Therefore for fixed $(x_2, \cdots, x_n, y_1, \cdots, y_p)$ near $(x_2^0, \cdots, x_n^0, y_1^0, \cdots, y_p^0)$, ϕ is analytic in x_1 for x_1 over $\mathfrak{W}_1 + L_1$, since the locus in question will be in $\mathfrak{E}$. Therefore the values determined by (3.1) are those of $\phi(x, y)$. Consequently ϕ is analytic at (x^0, y^0). The theorem follows immediately.

THEOREM 3.II. *Let* $\mathfrak{R}$ *be a limited domain, with connected boundary* B, *in the* $(2n + 1)$*-space of the real variable* y *and the complex variables* $x_1, \cdots, x_n$, $n > 0$, *and* $f(x_1, \cdots, x_n, y)$ *a function single-valued and analytic in an open set* $\mathfrak{S}$ *of the* $(2n + 2)$*-space of the complex variables* $x_1, \cdots, x_n, y$, *containing* B. *Then* f *admits a single-valued analytic continuation throughout* $\mathfrak{R} + \mathfrak{D}$, *where* $\mathfrak{D}$ *is the part of* $\mathfrak{S}$ *for which* y *is real.*

This means, of course, that f has an analytic continuation over a domain in $(2n + 2)$-space containing $\mathfrak{R} + \mathfrak{D}$. We may for the most part restrict ourselves to real values of y. The proof follows.

Let (α) denote a set of axes in the $(2n + 1)$-space, none of which is perpendicular to the y-axis. We divide $(2n + 1)$-space into $(2n + 1)$-cubes with edges parallel to the axes of the (α) system. As in the proof of Theorem 2.I, we let A be the locus of all of the closed cubes having points on $R = \mathfrak{R} + B$. Under definitions like those of that earlier case, C, part of the boundary of A, is the

connected locus of a $2n$-cycle, with each point of C accessible from infinity by a curve not meeting C, and $\mathfrak{T}$ is the limited domain bounded by C, with R a subset of $\mathfrak{T}$. We let T_0 denote $\mathfrak{T} + C$, and $\mathfrak{E} = T_0 + \mathfrak{D} = \mathfrak{T} + \mathfrak{D}$.

REMARK. *If a point Q of intersection of a plane Σ: y = real constant with C is not an isolated point of intersection, then C contains points near Q on both sides of Σ.* This is an easy consequence of the fact that the y-axis is not perpendicular to any axis of the (α) system.

If continuation over $\mathfrak{E}$ were impossible, there would exist a smallest value, e, of y, such that there would be a single-valued analytic function $\phi(x, y)$ defined over the part of $\mathfrak{E}$ for which $y > e$, and with $\phi(x, y) = f(x, y)$ in the part of $\mathfrak{D}$ not in T_0 for which $y > e$. Let Σ_0 be the plane $y = e$.

Let Q be any point of $\Sigma_0 \cdot \mathfrak{T}$. Plane Σ_0 intersects $\mathfrak{T}$ in a finite number of regions on Σ_0, and we let $\mathfrak{U}$ designate that one of those regions which contains Q. Then $\mathfrak{U}$ determines a $2n$-chain which is bounded (mod 2) by a $(2n - 1)$-cycle whose locus H is not necessarily connected. Let F denote the part of H consisting of all points of H accessible from infinity by curves on Σ_0 not meeting H. Then, as in similar situations arising above, F is also the locus of a $(2n - 1)$-cycle, bounding a limited domain $\mathfrak{F}$ on Σ_0 containing $\mathfrak{U}$; and F is connected. We infer from the remark above that near any point Z on F there are points of C above Σ_0 (where $y > e$). Therefore $\phi(x, y) = f(x, y)$ above Σ_0 near F.

For a moment we consider separately the cases $n > 1$ and $n = 1$.

If $n > 1$, we consider the analytic continuations of $f(x, y)$ over parts of the planes $y = e + \eta$ in (x, y)-space, with η any complex number near zero. For each η we continue $f(x, y)$ over the part of the plane which projects onto F and $\mathfrak{F}$, using Theorem 2.I. Then the continued function $\psi(x, y)$ is defined in the $(2n + 2)$-space in a neighborhood of Q among other points. Since it is easily seen that the hypotheses of Theorem 3.I are satisfied, where y is the parameter and B and $\mathfrak{R}$ of Theorem 3.I are taken as F and $\mathfrak{F}$, respectively, of the present proof, it follows that $\psi(x, y)$ is analytic near Q. Let us join Q by a curve l on $\mathfrak{U}$ to a point Z on F. Then f, ϕ and ψ are defined and all equal at and near the points of C near Z where $y > e$. The analytic continuation of ψ along and near l from Z to Q must equal ϕ in the part above Σ_0 of a region containing l. Therefore ϕ is analytically continuable throughout a neighborhood of Q, by means of the function $\psi(x, y)$.

If $n = 1$, let $\mathfrak{V}$ be the set obtained by adding to $\mathfrak{U}$ all its isolated boundary points Y, so that H is the boundary of $\mathfrak{V}$. Let H_1 and $\mathfrak{V}_1$ be the projections of H and $\mathfrak{V}$ respectively, on the x_1-plane, and $\psi(x_1, y) = \dfrac{1}{2\pi i}\displaystyle\int_{H_1} \frac{f(t, y)}{t - x_1}\,dt$, an analytic function near Q. But for real $y > e$ and $y - e$ small, $f(t, y) = \phi(t, y)$, and ϕ is analytic for x_1 on H_1 or $\mathfrak{V}_1$. Hence for x_1 in $\mathfrak{V}_1$ and $y > e$ with $y - e$ small, $\psi = \phi$. Thus in this case also we obtain a proper analytic continuation of ϕ in a neighborhood of Q, by means of the function[11] $\psi(x, y)$.

Case I. *Whenever a point Q is on $\Sigma_0 \cdot C$ and there are points of T_0 above Σ_0 near*

[11] This paragraph uses in part the method of Severi, loc. cit.

$Q, f(x, y) = \phi(x, y)$ *above* Σ_0 *near* Q. From the proof above it then follows that in Case I we could continue analytically throughout a neighborhood in $(2n + 2)$-space of each point Q of $\Sigma_0 \cdot T_0$, always getting proper values near C. It then follows easily that the number e used in the definition of Σ_0 could be replaced by a smaller number if Case I held, a contradiction.

Case II. *Not Case* I. There is at least one point Q_0 on C such that $f(x, y)$ neighboring Q_0 does not equal $\phi(x, y)$, where the latter is defined in that neighborhood. None of the part of C near Q_0 is above the plane $y = e$, for if this were the case, the values of $f(x, y)$ near Q_0 would have to equal those of $\phi(x, y)$ at and near the nearby points of C above Σ_0, and by the uniqueness of analytic continuation $f(x, y)$ would provide a continuation of $\phi(x, y)$ at Q_0 which would satisfy the conditions of Case I, contrary to hypothesis. Hence Q_0 is isolated as an intersection of Σ_0 and C, and y has a maximum at Q_0 on C, since if this were not the case it would follow from the remark above that there would be nearby points on C above the plane Σ_0.

From this point on the proof runs exactly like that of Case II in the proof of Theorem 2.I, with the planes $y =$ constant taking the place of the spheres with center at P in that proof, and with the remark above used occasionally. We need not repeat the details. It follows that Theorem 3.II is true.

4. Analytic homeomorphisms.

We begin with a topological property which is used in the final theorems.

THEOREM 4.I. *Let a self-compact point set A in a Hausdorff space be mapped in single-valued and continuous fashion on a set B in a Hausdorff space. If (i) the map is a homeomorphism for a neighborhood $\mathfrak{N}$ on A of each point P of A and the image of $\mathfrak{N}$; (ii) this image always forms a neighborhood on B of the image of P; and (iii) each point of B is either the image of only one point of A or can be joined by a curve*[12] *on B to such a point; then the map sets up a homeomorphism between A and B.*

The proof is similar to that in the more familiar case where B is assumed to be simply-connected.

DEFINITION. A homeomorphism set up between a locus A in $(\xi_1, \cdots, \xi_m)$-space and a locus A' in $(\xi_1', \cdots, \xi_m')$-space will be called *analytic* if defined by relations $\xi_j' = f_j(\xi_1, \cdots, \xi_m)$ $(j = 1, \cdots, m)$ where the functions are analytic in a domain containing A and the jacobian does not vanish on A. This definition is used both when the ξ's are real and when they are allowed to be complex.

THEOREM 4.II. *Let $\mathfrak{R}$ be a limited domain in the $2n$-space of the complex variables $x_1, \cdots, x_n$, $n > 1$, with connected boundary C. Let the equations $w_k = f_k(x_1, \cdots, x_n) = f_k(x)$ $(k = 1, \cdots, n)$ set up an analytic homeomorphism between C and a locus C' in (w)-space. The analytic continuations over $\mathfrak{R}$ of the functions $f_k(x)$ determine an analytic homeomorphism between $\mathfrak{R} + C$ and the image in (w)-space.*

Let $R = \mathfrak{R} + C$. We shall use Theorem 4.I, with $B = R'$, image of $R = A$. According to Theorem 2.I, the functions f_k do admit single-valued analytic

[12] A continuous image of a line segment is meant.

continuations over R. Since the jacobian J is not zero near C, its reciprocal is analytic there, and hence can be continued analytically over $\mathfrak{R}$. Therefore $J \neq 0$ in $\mathfrak{R}$ and it follows that (*i*) of Theorem 4.I is satisfied. Next we consider (*iii*).

Let Q be any point of R'. Draw a half-line (ray) through Q in any direction, and let Q_1, possibly Q itself, be the nearest point to Q on the half-line which is not an inner point on the whole line, say L, of the intersection set of L with R'. Now no point of $\mathfrak{R}$ can map on Q_1, since the fact that $J \neq 0$ would imply that R' contained a neighborhood of Q_1, from which a contradiction to the choice of Q_1 would follow. Hence Q_1 must be the image of a point of C, hence itself on C'. Since C and C' are in one-to-one correspondence it follows that Q_1 is the image of only one point of R. Hence (*iii*) of Theorem 4.I is satisfied.

As for (*ii*), this follows from the fact that $J \neq 0$, if P is a point of $\mathfrak{R}$. It will also follow when P is a point of C provided that we show that points outside a neighborhood $\mathfrak{N}$ of P are imaged on points outside some neighborhood of its image, say P'. But if this were not the case, we could find a limit point Q on $R - \mathfrak{N}$ having P' as image, contrary to the fact already proved that a point on C' is the image of only one point of R. Thus all the hypotheses of Theorem 4.I are satisfied, and it is seen that Theorem 4.II is true.

A closed connected non-vacuous locus S (a continuum) in the space of the real variables $\xi_1, \cdots, \xi_m$ will be called a *regular analytic* $(m-1)$-*spread* if neighboring each point of S it coincides with the locus of an equation of the form $\phi(\xi_1, \cdots, \xi_m) = 0$, where ϕ is real and analytic near P and $\phi_{\xi_1}^2 + \cdots + \phi_{\xi_m}^2 \neq 0$ there.

THEOREM 4.III. *Let* $\mathfrak{R}$ *be a limited domain in the* $(2n+1)$-*space of the real variable* y *and the complex variables* $x_1, \cdots, x_n$, $n > 0$, *with connected boundary* C. *Let* $\mathfrak{S}$ *be a domain in complex* (x, y)-*space containing* C, *and* $\mathfrak{D}$ *be the part of* $\mathfrak{S}$ *for which* y *is real. Suppose the equations* $w_j = f_j(x_1, \cdots, x_n, y)$ $(j = 1, 2, \cdots, n+1)$ *set up an analytic homeomorphism between* C *and a set* C', *and transform* $\mathfrak{D}$ *into a set on a regular analytic* $(2n+1)$-*spread* M. *Let* R' *denote the transform of* $R = \mathfrak{R} + C$ *under the analytic continuations over* $\mathfrak{R}$ *of the functions* $f_j(x, y)$.

Then R' *lies on* M. *Further, if* R' *does not cover all of* M, R' *is homeomorphic with* R.

For example, M might be a sphere. If M is a plane, then the final hypothesis is necessarily satisfied and hence need not be imposed.

We first prove that, under the analytic continuations over $\mathfrak{R}$ of the f's, which by Theorem 3.II exist, the image of R is entirely on M. By hypothesis this is true for points near C. Now suppose there were a point Q_1 on R whose image is not on M. We join Q_1 to a point Q_2 of C by a line on R. Let Q_3 be the point on that line farthest along from Q_2 towards Q_1 such that each point of the part of the line on the side of Q_3 towards Q_2 has a neighborhood on R all of whose points are mapped onto points of M by the transformation. Since M is closed, the image Q_3' of Q_3 is on M. Since M is a regular $(2n+1)$-spread, neighboring Q_3' it is the locus of an equation $\phi(u_1, v_1, \cdots, u_{n+1}, v_{n+1}) = 0$, with ϕ real and analytic, where $w_j = u_j + iv_j$ $(j = 1, \cdots, n+1)$. Now the equations of the transformation give the u's and the v's as real analytic functions of $r_1, t_1, \cdots, r_n,$

t_n, r_{n+1}, t_{n+1}, where $x_j = r_j + it_j$ $(j = 1, \cdots, n)$ and $y = r_{n+1} + it_{n+1}$, so that ϕ equals a function analytic in $r_1, \cdots, t_{n+1}$ for $(r_1, \cdots, t_{n+1})$ near Q_3. Then $\phi = 0$ at each point of an open set of a neighborhood of the projection of Q_3 on the space of $r_1, t_1, \cdots, r_n, t_n, r_{n+1}$ and hence is zero in the entire neighborhood. Therefore Q_3 cannot be the point farthest along the line from Q_2 satisfying the condition stated. We conclude that the image of R is on M.

The rest of the proof is similar to that of Theorem 4.II, aside from some considerations which we now mention.

It is easily shown that any connected part of M is connected by curves, by using the fact that any point of M has a neighborhood on M which is a $(2n + 1)$-cell. The "half-line through Q" of the proof of Theorem 4.II is replaced here by any curve through Q on M joining Q to a point of M which is not in R'. In proving that hypotheses (*iii*) and (*ii*) of Theorem 4.I are satisfied, we use the Brouwer theorem of invariance of regionality,[13] for the dimension $2n + 1$.

Columbia University.

[13] See Lefschetz I, page 100. Use of the theorem of invariance of regionality could be avoided by an argument involving jacobians.

TRANSFORMATIONS OF MULTIPLE SEQUENCES

By Hugh J. Hamilton

§1. Introduction and definition of notation

1.1. Notation. In order to treat n-tuple sequences with any degree of facility, it is necessary to introduce an abbreviated notation. The present paper uses one defined as follows.

The single letter m will denote an ordered set of n positive, integral variables, and k another such, homologous to m. A fixed value-set for m will be denoted by r, and the i-th of an infinite sequence of such sets by m_i. The symbols p and k_i are to be interpreted in an analogous sense with respect to k.

Generic representation for conjugate, proper, ordered subsets of any of these sets is to be obtained by affixing the superscripts 1 and 2, respectively, to the symbol denoting the set, and further subsets of like character with respect to either of these are to be represented by adjoining to the present superscript further numbers 1 and 2, respectively, etc. Two sets whose symbolic representations involve the same superscript are to be considered homologous. When the implication of this homology is *not* intended, the numbers 1 or 2 are replaced by 3 or 4, respectively, in one of the symbols. Thus k^3, k^4 are conjugate, but homologously independent of m^1, m^2.

A single element of k will be denoted generically by κ (or λ), and a fixed value of it by π. The corresponding element of m will be represented by μ.

All other letters are to be interpreted in the customary sense.

By relations like $k^{12} = p^{12}$ or $m_i > m_{i-1}$ are to be understood all sets of relations of the same form between corresponding elements of the two sets. In particular, the equation $k = p(m)$ is equivalent to the set of n equations $\kappa = \pi(m)$. The notations $k^3 = 1, 2, \cdots$ or $m^1 \geqq M$ imply the corresponding range of variation for each separate element of the set. However, inequalities like $k \nleqq M$ mean simply that *not every element* in the set is less than or equal to M.

Except when, by the nature of the situation, such would obviously be absurd, all relations involving subsets of k or of m are to be understood as implying the set of such relations for all possible choices of such subsets (with respect to position and, except when the subset consists only of κ or μ, with respect to dimension).

Received July 8, 1935. Presented to the American Mathematical Society, September 13, 1935. I am indebted to Professor C. R. Adams for suggesting the problems considered in this paper. A problem bearing a close analogy to the one considered here is treated in an article entitled *On transformations of double series*, which is expected to appear in the February, 1936 number of the Bulletin of the American Mathematical Society. Reference to this article may prove helpful to one who wishes to read carefully the present paper.

1.2. Principal types of sequences considered. The sequence $\{s_k\}$ is said to be *ultimately bounded*[1] (abbreviated ub) if there exists a number Q such that s_k is bounded for all $k > Q$; *bounded* (b) if in the preceding case Q can be taken to be zero; *convergent* (c) if $\lim_{k\to\infty} s_k \equiv s$ (read "principal limit") exists finite; *bounded convergent* (bc) if both b and c; *ultimately regularly convergent* (urc) if c, and if there exists a number $\bar{Q}$ such that $\lim_{k^1\to\infty} s_k \equiv s_{k^2}$ (read "row limit") exists finite for all $k^2 > \bar{Q}$; *regularly convergent*[2] (rc) if in the preceding case $\bar{Q}$ can be taken to be zero; *bounded ultimately regularly convergent* (burc) if both b and urc.

1.3. Nature of the transformation. A matrix $\| a_{mk} \|$ $(m, k = 1, 2, \cdots)$ is to be considered. By means of it the sequence $\{s_k\}$ is transformed into the sequence $\{\sigma_m\}$, where $\sigma_m \equiv \sum_{k=1}^{\infty} a_{mk} s_k$. Such a transformation is said to be *of infinite reference*, and the matrix *square*. If $a_{mk} = 0$ for all $\kappa > \pi(m)$ $(m = 1, 2, \cdots)$, then $\sigma_m = \sum_{k=1}^{p(m)} a_{mk} s_k$, and the transformation is said to be *of finite reference*, and the matrix *row finite*. If, in particular, $\pi(m) = \mu$, then $\sigma_m = \sum_{k=1}^{m} a_{mk} s_k$, and the matrix is said to be *triangular*.

1.4. Problems suggested. (i) If $\{s_k\}$ is of a specified one of the types defined in §1.2, under what conditions on $\| a_{mk} \|$ will $\{\sigma_m\}$ be of a specified one of these types? (ii) If $\{s_k\}$ is of a specified one of the last five types (involving convergence), under what conditions will $\{\sigma_m\}$ be of a specified one of these types, with $\sigma = s$? (The corresponding transformations and their matrices will be called *regular*.) (iii) If $\{s_k\}$ is of a specified one of the last three types (involving regular convergence, either complete or deferred), under what conditions will $\{\sigma_m\}$ be of a specified one of these types, with $\sigma_{k^2} = s_{k^2}$ for all k^2 sufficiently large? (The corresponding transformations and their matrices will be called *ultimately row regular*.) (iv) If $\{s_k\}$ is rc, under what conditions will $\{\sigma_m\}$ be rc with $\sigma_{k^2} = s_{k^2}$ for all k^2? (The corresponding transformation and its matrix will be called *row regular*.)

1.5. Auxiliary types of sequences. In order to attack these problems it is convenient to define under the classes of sequences listed in §1.3 several special types. Thus under the last five classes are introduced the corresponding *null sequences* $(s = 0)$, abbreviated cn, bcn, urcn, rcn, burcn, respectively; under the last three classes, the corresponding *ultimately row null sequences* $(s_{k^2} = 0$ for all k^2 greater than some number $\bar{Q})$, abbreviated urcrn, rcurn, burcrn,

[1] R. P. Agnew, American Journal of Mathematics, vol. 54 (1932), p. 648.

[2] G. H. Hardy, Proceedings of the Cambridge Philosophical Society, vol. 19 (1916–1919), p. 88.

respectively; under the rc type, *row null sequences* ($s_{k^2} = 0$ for all k^2), abbreviated rcrn.

1.6. **Notation for the transformation.** The process of transforming will be indicated by an arrow. "Necessary" will be abbreviated by N., "sufficient" by S., and "regularly", as applied to transformations (see 1.4 above), by reg. Thus S.c → rc reg reads "a set of conditions sufficient that every convergent sequence be transformed into a regularly convergent sequence with preservation of the principal limit".

1.7. **Existence of the transform.** Since under the infinite reference transformation σ_m may not even exist for certain values of m, it becomes necessary to speak of the class of sequences, *all of whose elements exist.* Such sequences are called *existent* (abbreviated e). In the present paper all proofs of the necessity of conditions are based on the assumption that $\{\sigma_m\}$ is e, though in the sufficiency proofs this requirement is not made (unless implied by the nature of the transformation).

1.8. **Earlier literature.** It appears that the questions proposed in (iii) and (iv) of §1.4 have thus far escaped attention, and that only a few of those in (i) and (ii) have been treated. Furthermore, of the four papers about to be cited, only one considers sequences of dimensionality greater than 2, and three contain slight errors in several of their conclusions. The nature of the results of these papers (in so far as they bear upon the problems suggested in §1.4) are indicated below. Unless otherwise stated, conditions will be lettered as in §3 of this paper.

Hallenbach[3] establishes conditions N. and S. for the transformations c → (e and c), b → b, bc → bc (square matrix, $n = 2$). The conditions are equivalent in each case to the corresponding conditions in the present paper.[4]

Kojima[5] finds conditions which he asserts to be N. and S. for cn → cn, cn → c, c → c, c → c reg, bc → c,[6] rc → c, rc → rc (triangular matrix, $n = 2$). In all but the last set of conditions, (c_1) should be replaced by (b_1), Kojima's proof of his inequality (6) being incorrect.[7] Also, the word "converges" in condition 3° of his last set[8] should be replaced by "converges regularly".

[3] Hallenbach, *Zur Theorie der Limitierungsverfahren von Doppelfolgen,* Dissertation, Bonn, 1933.

[4] Hallenbach, loc. cit. His conditions (B), (C) (first part) and (D) on p. 12 are equivalent to conditions (b_1) and (d_4) of this paper.

[5] Kojima, *On the theory of double sequences,* Tôhoku Mathematical Journal, vol. 21 (1922), pp. 3–14.

[6] Kojima, loc. cit., p. 12, Theorem V. Obvious typographical errors in the list of conditions can be remedied by studying the context.

[7] Kojima, loc. cit., p. 5. The choice indicated in lines 6 and 7 is not necessarily possible.

[8] Kojima, loc. cit., p. 14.

Leja[9] asserts conditions N. and S. for bc → c reg, bc → c, c → c reg, c → c (triangular matrix, $n = 2$). He then states conditions N. and S. for the same transformations (square matrix, $n = 2$), under the assumption that $\| a_{mk} \|$ satisfies (a_1) in all cases, and (a_2) in the last two. (As shown below, these conditions are actually N. that $\{\sigma_m\}$ be e.) Finally, he states with partial proof conditions N. and S. for the same transformations (triangular matrix, $n = n$), under the assumption in the last two cases that the matrix satisfies an added condition equivalent to (b_2), which is proved to be N. in §5 below. In all cases, (c_1) should be replaced by (b_1), since Leja's proof of the necessity of his condition 3° is in error.[10]

Robison[11] asserts conditions N. and S. for bc → c reg (triangular matrix, $n = 2$); bc → (e and c reg) (square matrix, $n = 2$). In both cases (c_1) should be replaced by (b_1), the proof of Robison's inequalities (4) being at fault.[12] He correctly states conditions N. and S. for bc → bc (triangular matrix, $n = 2$); bc → (e and bc) (square matrix, $n = 2$). Conditions are also stated to be N. and S. for bc → c, with σ a function of s only (triangular matrix, $n = 2$); and for bc → (e and c), with σ a function of s only (square matrix, $n = 2$): in the first set (c_1) should be replaced by (b_1) and in the other (b_1) should be added. Finally, conditions are stated to be N. and S. for b → bc (triangular matrix, $n = 2$); b → (e and bc) (square matrix, $n = 2$), both sets being incorrect: the first can be rectified by replacing Robison's second condition by (c_1), and the second by replacing his first condition by (c_1) and omitting his last.

1.9. **Scope of the present paper.** In this paper are found conditions N. that a sequence of any specified one of the types listed in §§1.2 and 1.5 have an existent transform which is of a specified one of these types, and conditions S. that the transform of such a sequence be of a specified one of these types, without being necessarily defined for all m. The hypothesis of existence of the transform is of course implied when the latter is bounded, and the manner of obtaining additional conditions to insure that the transform exist in any case will appear.

In the sense of the preceding paragraph, then, the question (i) of §1.4 is completely answered. The answer to (ii) is given in §7. For the special cases of ultimately row null and row null sequences, questions (iii) and (iv) are answered, though the complete solutions seem at present to offer difficulties.

§2 below contains certain preliminary lemmas and some remarks on the proofs to follow. In §3 are listed the conditions to be used, various implications of which are indicated in §4. N. proofs are given in §5, and S. proofs in §6. A

[9] Leja, *Sur les transformations linéaires des suites doubles et multiples*, Bulletin International de l'Académie Polonaise, Classe des Sciences, A, (1930), pp. 1–10.

[10] Leja, loc. cit., p. 4. The statement in lines 6 and 7 is untrue.

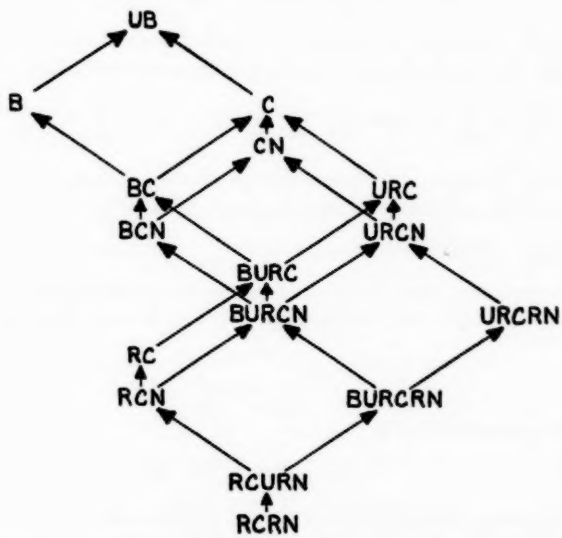
[11] Robison, *Divergent double sequences and series*, Transactions of the American Mathematical Society, vol. 28 (1926), pp. 50–73.

[12] Robison, loc. cit., p. 55, line 13. Cf. footnote 7.

few remarks on interpretation of results and a list of conditions for regularity in §7 conclude the paper.

§2. Preliminary observations

Under the hypothesis that rc sequences are e, the several types of sequences to be considered are related as shown in the diagram, the arrow indicating implication of the quality at its head by that at its tail.



These relations are sufficiently clear, save perhaps that regular convergence plus existence implies boundedness, but this is an immediate corollary of the theorem below on uniformity of row convergence.

Tacit use is made throughout of the fact that, if u $\subset$ U and v $\subset$ V, where u, U, v, V represent classes of sequences, then N.u $\to$ V is at once N.(u, U) $\to$ (v, V) (four cases), and S.U $\to$ v is at once S.(u, U) $\to$ (v, V).

In accordance with the remark in the last paragraph of §1.1, it is to be noted that most of the conditions in §3 are equivalent to sets of several conditions. Thus (f_4) implies $n(2^n - 2)$ separate conditions depending on the various possible positions of κ, the dimensionalities of m^2 (1 to $n - 1$), and the various ways in which m^2 can be chosen positionally from m for fixed dimensionality. However, in necessity proofs, where denial of such a condition implies denial for a particular positionality of κ and a particular positionality and dimensionality of m^2, these qualities are necessarily assumed fixed throughout any given theorem. If certain variables are given particular values in the course of a proof, this fact will be indicated in parentheses (or brackets), and subsequently parentheses (or

brackets) will enclose the generic symbol involved. Thus: "With $(m^1 = r^1)$ and $(k^3 = p^3)$, $\lim_{m^2 \to \infty} \sum_{k^4=1}^{\infty} a_{(m)(k)} = L_{r^1 p^3}$".

It is often convenient to decompose the operator $\sum_{k=1}^{\infty}$ as follows (R being an arbitrary positive integer):

$$\sum_{k=1}^{\infty} = \sum \sum_{k^1=1}^{R} \sum_{k^2=R+1}^{\infty} + \sum_{k=1}^{R} + \sum_{k=R+1}^{\infty}. \tag{.001}$$

Here the free $\sum$ indicates summation over all manners of choice of k^1 from k, both dimensionally and positionally. Another decomposition is the following:

$$\sum_{k=1}^{\infty} = \sum_{\nu=1}^{n-1} (-1)^{\nu+1} \sum \sum_{k^1=1}^{R} \sum_{k^2=1}^{\infty} + (-1)^{n+1} \sum_{k=1}^{R} + \sum_{k=R+1}^{\infty}, \tag{.002}$$

where the free $\sum$ sums over all manners of positional choice of the summation index immediately following, for dimensionality defined by the summation index first preceding. Similar decompositions of $\sum_{k^2=1}^{\infty}$ occur, etc.

A theorem on uniformity of row convergence of rc sequences is now given.

(.003) THEOREM. *If s_k is rc, then, given any $\epsilon > 0$, there exists $R \equiv R(\epsilon)$ such that*

$$| s_k - s_{k^1} | < 2\epsilon \text{ for } k^2 > R \qquad (k^1 = 1, 2, \cdots) \tag{.0001}$$

From the convergence of s_k follows

$$| s_k - s | < \epsilon \text{ for } k > R_1. \tag{.0002}$$

Since s_k is rc, there exist $R_1 < R_2 < \cdots < R_n$ such that

$$| s_k - s_{k^1} | < \epsilon \text{ for } k^2 > R_{i+1} \qquad (k^1 = 1, 2, \cdots, R_i;\ i = 1, 2, \cdots, n-1). \tag{.0003}$$

Let p^1 be arbitrary. Since the dimensionality of p^1 is not greater than $n - 1$, at least one interval $(R_i + 0, R_{i+1})$ $(i = 0, 1, \cdots, n - 1;\ R_0 \equiv 0)$ contains no element of it. If $p^1 \leqq R_i$, from (.0003) it follows, with $(k^1 = p^1)$, that $| s_{(k)} - s_{p^1} | < \epsilon$ for $k^2 > R_{i+1}$; if $p^1 > R_{i+1}$, from (.0002) it follows that $| s_{(k)} - s_{p^1} | < 2\epsilon$ for $k^2 > R_1$; if $p^{11} \leqq R_i$ and $p^{12} > R_{i+1}$, from (.0003) it follows that $| s_{(k)} - s_{p^1} | \leqq | s_{(k)} - s_{p^{11}} | + | s_{p^1} - s_{p^{11}} | < 2\epsilon$ for $k^2 > R_{i+1}$. Hence $R \equiv R_n$ satisfies (.0001).

In the course of the proofs in §5 it becomes convenient to use the *oscillation at infinity* of a function $f(m)$, defined thus: $\operatorname{osc}_{m=\infty} f(m) \equiv \overline{\lim}_{m,r \to \infty} | f(m) - f(r) |$. At once it follows that $\operatorname{osc}_{m=\infty} f(m) - \operatorname{osc}_{m=\infty} \phi(m) \leqq \operatorname{osc}_{m=\infty} \{f(m) + \phi(m)\} \leqq \operatorname{osc}_{m=\infty} f(m) + \operatorname{osc}_{m=\infty} \phi(m)$; $\operatorname{osc}_{m=\infty} f(m) \leqq 2 \overline{\lim}_{m \to \infty} | f(m) |$; $\operatorname{osc}_{m=\infty} C \cdot f(m) = C \cdot \operatorname{osc}_{m=\infty} f(m)$; and N. and S. $\lim_{m \to \infty} f(m)$ exist finite is: $\operatorname{osc}_{m=\infty} f(m) = 0$.

Use is made in §6 of the following set of sufficient conditions that $\lim\limits_{m\to\infty} f(m) = \phi$, namely: $f(m) \equiv \phi(m, R) + \psi(m, R)$, where R is an arbitrary positive integer; $\lim\limits_{m,R\to\infty} \psi(m, R) = 0$; $\lim\limits_{m\to\infty} \phi(m, R) = \phi(R)$ for each R; $\lim\limits_{R\to\infty} \phi(R) = \phi$.

§3. Conditions on the matrix

Existence conditions.

(a₁) $$\sum_{k=1}^{\infty} |a_{mk}| < \infty \qquad (m = 1, 2, \cdots).$$

(a₂) Let κ, λ be any two single elements of k, and k^2 those remaining. Then

$$a_{mk} = 0 \text{ for } \lambda > C_\kappa(m) \qquad (k^2 = 1, 2, \cdots; m = 1, 2, \cdots; \kappa = 1, 2, \cdots).$$

UB conditions.

(b₁) $$\sum_{k=1}^{\infty} |a_{mk}| < A \text{ for } m > B.$$

(b₂) Let κ, λ be any two single elements of k, and k^2 those remaining. Then

$$a_{mk} = 0 \text{ for } m, \lambda > C_\kappa \qquad (k^2 = 1, 2, \cdots; \kappa = 1, 2, \cdots).$$

B conditions.

(c₁) $$\sum_{k=1}^{\infty} |a_{mk}| < A \qquad (m = 1, 2, \cdots).$$

(c₂) Let κ, λ be any two single elements of k, and k^2 those remaining. Then

$$a_{mk} = 0 \text{ for } \lambda > C_\kappa \qquad (k^2 = 1, 2, \cdots; m = 1, 2, \cdots; \kappa = 1, 2, \cdots).$$

C conditions.

(d₁) $$\lim_{m\to\infty} a_{mk} = a_k \qquad (k = 1, 2, \cdots).$$

(d₂) $$\lim_{m\to\infty} \sum_{k^2=1}^{\infty} a_{mk} = L_{k^1} \qquad (k^1 = 1, 2, \cdots).$$

(d₃) $$\lim_{m\to\infty} \sum_{k=1}^{\infty} a_{mk} = L.$$

(d₄) There exist numbers a_k such that, if κ is any single element of k, and k^2 those remaining, then

$$\lim_{m\to\infty} \sum_{k^2=1}^{\infty} |a_{mk} - a_k| = 0 \qquad (\kappa = 1, 2, \cdots).$$

(d₅) There exist numbers a_k such that

$$\lim_{m\to\infty} \sum_{k=1}^{\infty} |a_{mk} - a_k| = 0.$$

CN conditions.

$(\bar{d}_1)$ $\qquad$ (d_1), with $a_k = 0$ $\qquad$ $(k = 1, 2, \cdots)$.

$(\bar{d}_2)$ $\qquad$ (d_2), with $L_{k^1} = 0$ $\qquad$ $(k^1 = 1, 2, \cdots)$.

$(\bar{d}_3)$ $\qquad$ (d_3), with $L = 0$.

$(\bar{d}_4)$ $\qquad$ (d_4), with $a_k = 0$ $\qquad$ $(k = 1, 2, \cdots)$.

$(\bar{d}_5)$ $\qquad$ (d_5), with $a_k = 0$ $\qquad$ $(k = 1, 2, \cdots)$.

URC conditions.

(e_1)
$$\lim_{m^2 \to \infty} a_{mk} = a_{m^1k} \text{ for } m^1 > D \qquad (k = 1, 2, \cdots).$$

(e_2^*) Let κ be any single element of k^3, and k^{32} (which may in this instance be null) those remaining. Then

$$\lim_{m^2 \to \infty} \sum_{k^4=1}^{\infty} a_{mk} = L_{m^1k^3} \text{ for } m^1 > E_\kappa \qquad (k^{32} = 1, 2, \cdots; \kappa = 1, 2, \cdots).$$

(e_2)
$$\lim_{m^2 \to \infty} \sum_{k^4=1}^{\infty} a_{mk} = L_{m^1k^3} \text{ for } m^1 > E \qquad (k^3 = 1, 2, \cdots).$$

(e_3)
$$\lim_{m^2 \to \infty} \sum_{k=1}^{\infty} a_{mk} = L_{m^1} \text{ for } m^1 > F.$$

(e_4^*) There exist numbers a_{m^1k} such that, if κ is any single element of k, and k^4 those remaining, then

$$\lim_{m^2 \to \infty} \sum_{k^4=1}^{\infty} | a_{mk} - a_{m^1k} | = 0 \text{ for } m^1 > G_\kappa \qquad (\kappa = 1, 2, \cdots).$$

(e_4) There exist numbers a_{m^1k} such that, if κ is any single element of k, and k^4 those remaining, then

$$\lim_{m^2 \to \infty} \sum_{k^4=1}^{\infty} | a_{mk} - a_{m^1k} | = 0 \text{ for } m^1 > G \qquad (\kappa = 1, 2, \cdots).$$

(e_5) There exist numbers a_{m^1k} such that

$$\lim_{m^2 \to \infty} \sum_{k=1}^{\infty} | a_{mk} - a_{m^1k} | = 0 \text{ for } m^1 > H.$$

URCRN conditions.

$(\bar{e}_1)$ $\qquad$ (e_1), with $a_{m^1k} = 0$ for $m^1 > \bar{D}$ $\qquad$ $(k = 1, 2, \cdots)$.

$(\bar{e}_2^*)$ $\qquad$ (e_2^*), with $L_{m^1k^3} = 0$ for $m^1 > \bar{E}_\kappa$

$(k^{32} = 1, 2, \cdots; \kappa = 1, 2, \cdots)$.

$(\bar{e}_2)$ $\qquad$ (e_2), with $L_{m^1k^3} = 0$ for $m^1 > \bar{E}$ $\qquad$ $(k^3 = 1, 2, \cdots)$.

$(\bar{e}_3)$ (e_3), with $L_{m^1} = 0$ for $m^1 > \bar{F}$.

$(\bar{e}_4^*)$ (e_4^*), with $a_{m^1k} = 0$ for $m^1 > \bar{G}_\kappa$

$$(k^4 = 1, 2, \cdots; \kappa = 1, 2, \cdots).$$

$(\bar{e}_4)$ (e_4), with $a_{m^1k} = 0$ for $m^1 > \bar{G}$ $(k = 1, 2, \cdots)$.

$(\bar{e}_5)$ (e_5), with $a_{m^1k} = 0$ for $m^1 > \bar{H}$ $(k = 1, 2, \cdots)$.

RC conditions.

(f_1) $$\lim_{m^2\to\infty} a_{mk} = a_{m^1k} \text{ for all } m^1 \qquad (k = 1, 2, \cdots).$$

(f_2) $$\lim_{m^2\to\infty} \sum_{k^4=1}^{\infty} a_{mk} = L_{m^1k^3} \text{ for all } m^1 \qquad (k^3 = 1, 2, \cdots).$$

(f_3) $$\lim_{m^2\to\infty} \sum_{k=1}^{\infty} a_{mk} = L_{m^1} \text{ for all } m^1.$$

(f_4) There exist numbers a_{m^1k} such that, if κ is any single element of k, and k^4 those remaining, then

$$\lim_{m^2\to\infty} \sum_{k^4=1}^{\infty} | a_{mk} - a_{m^1k} | = 0 \text{ for all } m^1 \qquad (\kappa = 1, 2, \cdots).$$

(f_5) There exist numbers a_{m^1k} such that

$$\lim_{m^2\to\infty} \sum_{k=1}^{\infty} | a_{mk} - a_{m^1k} | = 0 \text{ for all } m^1.$$

RCRN conditions.

$(\mathring{f}_1)$ (f_1), with $a_{m^1k} = 0$ for all m^1 $(k = 1, 2, \cdots)$.

$(\mathring{f}_2)$ (f_2), with $L_{m^1k^3} = 0$ for all m^1 $(k^3 = 1, 2, \cdots)$.

$(\mathring{f}_3)$ (f_3), with $L_{m^1} = 0$ for all m^1.

$(\mathring{f}_4)$ (f_4), with $a_{m^1k} = 0$ for all m^1 $(k = 1, 2, \cdots)$.

$(\mathring{f}_5)$ (f_5), with $a_{m^1k} = 0$ for all m^1 $(k = 1, 2, \cdots)$.

§4. Implications of the conditions

For later use it is convenient to list the following relations, which may be easily established. The symbol $\to$ indicates implication.

(.01) $$(b_1) + (d_1) \to \sum_{k=1}^{\infty} | a_k | \leqq A.$$

(.02) $$(b_1) + (d_2) \to \sum_{k^1=1}^{\infty} | L_{k^1} | \leqq A.$$

(.03) $(b_1) + (e_1) \rightarrow \sum_{k=1}^{\infty} | a_{m^1k} | \leqq A$ for $m^1 > B, D$.

(.04) $(b_1) + (e_2^*) \rightarrow \sum_{k^3=1}^{\infty} | L_{m^1k^3} | \leqq A$ for $m^1 > B, E_\kappa$ $(\kappa = 1, 2, \cdots)$

(.05) $(b_1) + (e_2) \rightarrow \sum_{k^3=1}^{\infty} | L_{m^1k^3} | \leqq A$ for $m^1 > B, E$.

(.06) $(c_1) + (f_1) \rightarrow \sum_{k=1}^{\infty} | a_{m^1k} | \leqq A$ for all m^1.

(.07) $(c_1) + (f_2) \rightarrow \sum_{k^3=1}^{\infty} | L_{m^1k^3} | \leqq A$ for all m^1.

(.08) $(c_1) \rightarrow (a_1)$.

(.09) $(c_1) \rightarrow (b_1)$, with $B = 0$.

(.10) $(c_2) \rightarrow (a_2)$, with $C_\kappa(m) = C_\kappa$ for all m $(\kappa = 1, 2, \cdots)$.

(.11) $(c_2) \rightarrow (b_2)$.

(.12) $(b_2) + (d_1) \rightarrow (d_4)$.

(.13) $(d_4) \rightarrow (d_1)$.

(.14) $(b_1) + (d_4) \rightarrow (d_2)$, with $L_{k^1} = \sum_{k^2=1}^{\infty} a_k$ $(k^1 = 1, 2, \cdots)$.

(.15) $(d_5) \rightarrow (d_4)$.

(.16) $(b_1) + (d_5) \rightarrow (d_3)$, with $L = \sum_{k=1}^{\infty} a_k$.

(.17) Each CN condition implies the corresponding C condition.

(.18) $(\bar{d}_1) + (d_4) \rightarrow (\bar{d}_4)$.

(.19) $(\bar{d}_1) + (d_5) \rightarrow (\bar{d}_5)$.

(.20) $(\bar{d}_4) \rightarrow (\bar{d}_1)$.

(.21) $(\bar{d}_4) \rightarrow (\bar{d}_2)$.

(.22) $(\bar{d}_5) \rightarrow (\bar{d}_3)$.

(.23) $(\bar{d}_5) \rightarrow (\bar{d}_4)$.

(.24) $(b_2) + (e_1) \rightarrow (e_4^*)$, with $G_\kappa = \max (C_\kappa, D)$ $(\kappa = 1, 2, \cdots)$.

(.25) $(c_2) + (e_1) \rightarrow (e_4)$, with $G = D$.

(.26) $(e_2) \rightarrow (e_2^*)$, with $E_\kappa = E$ $(\kappa = 1, 2, \cdots)$.

(.27) $(\mathrm{b}_1) + (\mathrm{e}_4^*) \rightarrow (\mathrm{e}_2^*)$, with $E_\kappa = \max(B, G_\kappa)$ $(\kappa = 1, 2, \cdots)$, and

$$L_{m^1k^3} = \sum_{k^4=1}^{\infty} a_{m^1k} \qquad (k^3 = 1, 2, \cdots).$$

(.28) $(\mathrm{e}_4) \rightarrow (\mathrm{e}_1)$, with $D = G$.

(.29) $(\mathrm{b}_1) + (\mathrm{e}_4) \rightarrow (\mathrm{e}_2)$, with $E = \max(B, G)$, and

$$L_{m^1k^3} = \sum_{k^4=1}^{\infty} a_{m^1k} \qquad (k^3 = 1, 2, \cdots).$$

(.30) $(\mathrm{e}_4) \rightarrow (\mathrm{e}_4^*)$, with $G_\kappa = G$ $(\kappa = 1, 2, \cdots)$.

(.31) $(\mathrm{e}_5) \rightarrow (\mathrm{e}_4)$, with $G = H$.

(.32) $(\mathrm{b}_1) + (\mathrm{e}_5) \rightarrow (\mathrm{e}_3)$, with $F = \max(B, H)$, and $L_{m^1} = \sum_{k=1}^{\infty} a_{m^1k}$.

(.33) Each URCRN condition implies the corresponding URC condition, with D, E, etc., replaced by $\bar{D}, \bar{E}$, etc., respectively.

(.34) $(\bar{\mathrm{e}}_1) + (\mathrm{e}_4^*) \rightarrow (\bar{\mathrm{e}}_4^*)$, with $\bar{G}_\kappa = \max(\bar{D}, G_\kappa)$ $(\kappa = 1, 2, \cdots)$.

(.35) $(\bar{\mathrm{e}}_1) + (\mathrm{e}_4) \rightarrow (\bar{\mathrm{e}}_4)$, with $\bar{G} = \max(\bar{D}, G)$.

(.36) $(\bar{\mathrm{e}}_1) + (\mathrm{e}_5) \rightarrow (\bar{\mathrm{e}}_5)$, with $\bar{H} = \max(\bar{D}, H)$.

(.37) $(\bar{\mathrm{e}}_2) \rightarrow (\bar{\mathrm{e}}_2^*)$, with $\bar{E}_\kappa = \bar{E}$ $(\kappa = 1, 2, \cdots)$.

(.38) $(\bar{\mathrm{e}}_4^*) \rightarrow (\bar{\mathrm{e}}_2^*)$, with $\bar{E}_\kappa = \bar{G}_\kappa$ $(\kappa = 1, 2, \cdots)$.

(.39) $(\bar{\mathrm{e}}_4) \rightarrow (\bar{\mathrm{e}}_1)$, with $\bar{D} = \bar{G}$.

(.40) $(\bar{\mathrm{e}}_4) \rightarrow (\bar{\mathrm{e}}_2)$, with $\bar{E} = \bar{G}$.

(.41) $(\bar{\mathrm{e}}_4) \rightarrow (\bar{\mathrm{e}}_4^*)$, with $\bar{G}_\kappa = \bar{G}$ $(\kappa = 1, 2, \cdots)$.

(.42) $(\bar{\mathrm{e}}_5) \rightarrow (\bar{\mathrm{e}}_3)$, with $\bar{F} = \bar{H}$.

(.43) $(\bar{\mathrm{e}}_5) \rightarrow (\bar{\mathrm{e}}_4)$, with $\bar{G} = \bar{H}$.

(.44) $(\mathrm{c}_2) + (\mathrm{f}_1) \rightarrow (\mathrm{f}_4)$.

(.45) $(\mathrm{f}_4) \rightarrow (\mathrm{f}_1)$.

(.46) $(\mathrm{c}_1) + (\mathrm{f}_4) \rightarrow (\mathrm{f}_2)$, with $L_{m^1k^3} = \sum_{k^4=1}^{\infty} a_{m^1k}$ $(k^3 = 1, 2, \cdots)$.

(.47) $(\mathrm{f}_5) \rightarrow (\mathrm{f}_4)$.

(.48) $(\mathrm{c}_1) + (\mathrm{f}_5) \rightarrow (\mathrm{f}_3)$, with $L_{m^1} = \sum_{k=1}^{\infty} a_{m^1k}$.

(.49) Each RC condition implies the corresponding URC condition for all m^1.

(.50) Each RCRN condition implies the corresponding RC condition.

(.51) $(\bar{f}_1) + (f_4) \rightarrow (\bar{f}_4)$.

(.52) $(\bar{f}_1) + (f_5) \rightarrow (\bar{f}_5)$.

(.53) $(\bar{f}_4) \rightarrow (\bar{f}_1)$.

(.54) $(\bar{f}_4) \rightarrow (\bar{f}_2)$.

(.55) $(\bar{f}_5) \rightarrow (\bar{f}_3)$.

(.56) $(\bar{f}_5) \rightarrow (\bar{f}_4)$.

(.57) Each RCRN condition implies the corresponding URCRN condition for all m^1.

(Brackets in the following conditions indicate application of the corresponding parenthesized condition to the matrix $\| a_{mk} \|'$, defined for all m, but only for $k > Q$, where Q is an arbitrary positive integer.)

(.58) With the exception of those in the right-hand members below, each condition implies its bracketed counterpart.

(.59) $(d_1) + (d_2) \rightarrow [d_2]$.

(.60) $(d_1) + (d_2) + (d_3) \rightarrow [d_3]$.

(.61) $(\bar{d}_1) + (\bar{d}_2) \rightarrow [\bar{d}_2]$.

(.62) $(\bar{d}_1) + (\bar{d}_2) + (\bar{d}_3) \rightarrow [\bar{d}_3]$.

(.63) $(e_1) + (e_2^*) \rightarrow [e_2^*]$ for $m^1 > \max(D, E_\kappa)$ $(\kappa = 1, 2, \cdots)$.

(.64) $(e_1) + (e_2) \rightarrow [e_2]$ for $m^1 > \max(D, E)$.

(.65) $(e_1) + (e_2) + (e_3) \rightarrow [e_3]$ for $m^1 > \max(D, E, F)$.

(.66) $(\bar{e}_1) + (\bar{e}_2^*) \rightarrow [\bar{e}_2^*]$ for $m^1 > \max(\bar{D}, \bar{E}_\kappa)$ $(\kappa = 1, 2, \cdots)$.

(.67) $(\bar{e}_1) + (\bar{e}_2) \rightarrow [\bar{e}_2]$ for $m^1 > \max(\bar{D}, \bar{E})$.

(.68) $(\bar{e}_1) + (\bar{e}_2) + (\bar{e}_3) \rightarrow [\bar{e}_3]$ for $m^1 > \max(\bar{D}, \bar{E}, \bar{F})$.

(.69) $(f_1) + (f_2) \rightarrow [f_2]$.

(.70) $(f_1) + (f_2) + (f_3) \rightarrow [f_3]$.

(.71) $(\bar{f}_1) + (\bar{f}_2) \rightarrow [\bar{f}_2]$.

(.72) $(\bar{f}_1) + (\bar{f}_2) + (\bar{f}_3) \rightarrow [\bar{f}_3]$.

(.73) $(d_1) + (\bar{e}_1) \rightarrow (\bar{d}_1)$.

(.74) $(d_2) + (\bar{e}_2^*) \rightarrow (\bar{d}_2)$.

(.75) $(d_3) + (\bar{e}_3) \to (\bar{d}_3)$.

(.76) $(\bar{d}_5) \to (b_1)$.

(.77) $(a_1) + (d_5) + (f_5) \to (c_1)$.

The methods of proof of these relations will be sufficiently clear in view of the following typical examples.

Proof of (.12). From (b_2) it follows that $a_k = 0$ for $\lambda > C_\kappa$ ($k^2 = 1, 2, \cdots$). Hence, for $m > C_\kappa$, in the notation of (d_4), $\sum_{k^2=1}^{\infty} |a_{mk} - a_k| = \sum_{k^2=1}^{C_\kappa} |a_{mk} - a_k|$.

Proof of (.59). By (.002), with the dimensionality of k^2 represented by t, it follows that

$$\sum_{k^2=Q+1}^{\infty} a_{mk} = \sum_{k^2=1}^{\infty} a_{mk} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{21}=1}^{Q} \sum_{k^{22}=1}^{\infty} a_{mk} - (-1)^{t+1} \sum_{k^2=1}^{Q} a_{mk}\,.$$

Proof of (.73). For fixed k, the sequence in m, $\{a_{mk}\}$, is urc, with all row-limits ultimately zero. Hence the principal limit is zero.

§5. Necessity proofs (See §1.7)

1. *N.RCRN* $\to$ *e is* (a_1). By denial of (a_1), there exists an r such that $\sum_{k=1}^{\infty} |a_{rk}| = \infty$. Let $M_i > M_{i-1}$ be such that $\sum_{k=1, \nleq M_{i-1}}^{M_i} |a_{rk}| \geqq i$. Now $s_k \equiv (-1)^i \operatorname{sgn} a_{rk}/i$ ($k \leqq M_i$, $\nleq M_{i-1}$) is rcrn, while σ_r does not exist.

2. *N.URCRN* $\to$ *e is* (a_2). By denial of (a_2), there exist π, r, and sequences k_i^2 and λ_i ($> \lambda_{i-1}$), such that, with $\kappa_i = \pi$ ($i = 1, 2, \cdots$), $a_{rk_i} \neq 0$. Now $s_k \equiv \{(-1)^i/a_{rk}$ ($k = k_i$); 0, otherwise$\}$ is urcrn, while $\sigma_r = -1 + 1 - 1 + \cdots$.

3. *N.RCRN* $\to$ *UB is* (b_1). Since the existence of the transform is assumed, (a_1) can be assumed, by 1. Now $|a_{mk}| < A_k$ for $m > B_k$ ($k = 1, 2, \cdots$). For, with arbitrary p, $s_k \equiv \{1$ ($k = p$); 0, otherwise$\}$ is rcrn, and $\sigma_m = a_{mp}$ must be ub.

By denial of (b_1), there exist sequences M_i, m_i, satisfying $M_i > M_{i-1}$;

$$m_i > m_{i-1}, B_k \qquad (k \leqq M_{i-1})\,, \tag{3.1}$$

such that $\sum_{k=1}^{M_i} |a_{m_ik}| \geqq i\left(2 \sum_{k=1}^{M_{i-1}} A_k + i\right)$, where, by (a_1), $\sum_{k=1, \nleq M_i}^{\infty} |a_{m_ik}| \leqq 1$. The sequence $s_k \equiv \operatorname{sgn} a_{m_ik}/i$ ($k \leqq M_i$, $\nleq M_{i-1}$) is rcrn, whereas

$$|\sigma_{m_i}| \geqq \sum_{k=1}^{M_i} |a_{m_ik}|/i - 2\sum_{k=1}^{M_{i-1}} A_k - \sum_{k=1, \nleq M_i}^{\infty} |a_{m_ik}| \geqq i - 1\,,$$

so that, by (3.1), σ_m is not ub.

4. *N.URCRN* $\to$ *UB is* (b_2). (a_2) can be assumed, by 2. By denial of (b_2), there exist π and sequences k_i^2 and λ_i, m_i, satisfying $\lambda_i > \lambda_{i-1}$, $C_\pi(m_j)$ ($j < i$),

$$m_i > m_{i-1}\,, \tag{4.1}$$

such that, with $\kappa_i = \pi$ $(i = 1, 2, \cdots)$, $a_{m_i k_i} \neq 0$. Now $s_k \equiv \left\{1 \ (k = k_1);\ \left(\sum_{j=1}^{i-1} | a_{m_i k_j} s_{k_j} | + i\right) \Big/ a_{m_i k_i} \ (k = k_i \text{ for } i > 1);\ 0, \text{ otherwise}\right\}$ is urcrn, while $| \sigma_{m_i} | = \left| \sum_{j=1}^{i-1} a_{m_i k_j} s_{k_j} + a_{m_i k_i} s_{k_i} \right| \geqq i$. Hence, by (4.1), σ_m is not ub.

5. *N.RCRN* $\rightarrow$ *B is* (c_1). Proof similar to that of 3.

6. *N.URCRN* $\rightarrow$ *B is* (c_2). Proof similar to that of 4.

7. *N.RCRN* $\rightarrow$ *C is* (d_1). Let p be arbitrary. Now $s_k \equiv \{1\ (k = p);\ 0, \text{otherwise}\}$ is rcrn, and $\sigma_m = a_{mp}$.

8. *N.RCURN* $\rightarrow$ *C is* (d_2). Let p^1 be arbitrary. The sequence $s_k \equiv \{1\ (k^1 = p^1;\ k^2 = 1, 2, \cdots);\ 0, \text{otherwise}\}$ is rcurn, and, with $(k^1 = p^1)$, $\sigma_m = \sum_{k^2=1}^{\infty} a_{m(k)}$.

9. *N.RC* $\rightarrow$ *C is* (d_3). The sequence $s_k = 1$ $(k = 1, 2, \cdots)$ is rc, while $\sigma_m = \sum_{k=1}^{\infty} a_{mk}$.

10. *N. BURCRN* $\rightarrow$ *C is* (d_4). (b_1) is assumed, by 3, and (d_1), by 7. By denial of (d_4), then, there exist $\delta > 0$, π, and sequences $M_i (> M_{i-1})$ and m_i satisfying

$$m_i > m_{i-1}, B, \tag{10.1}$$

such that, with $(\kappa = \pi)$, $\sum_{k^2=1}^{M_i} | a_{m_i(k)} - a_{(k)} | \geqq 6\delta$, where, by (10.1), ($b_1$), and (.01), $\sum_{k^2=1,\ \nleqq M_i}^{\infty} | a_{mi(k)} - a_{(k)} | \leqq \delta$, and, by ($d_1$), $\sum_{k^2=1}^{M_{i-1}} | a_{m_i(k)} - a_{(k)} | \leqq \delta$.

The sequence $s_k \equiv \{\text{sgn}(a_{m_i k} - a_k)\ (\kappa = \pi;\ k^2 \leqq M_i, \nleqq M_{i-1}, \text{ for } i \text{ odd});\ 0, \text{otherwise}\}$ is burcrn. Now

$$\begin{aligned} \sigma_{m_i} &= \sum_{k^2=1}^{\infty} a_{(k)} s_{(k)} && \text{(A)} \\ &+ \sum_{k^2=1}^{\infty} (a_{m_i(k)} - a_{(k)}) s_{(k)} . && \text{(B)} \end{aligned} \tag{10.2}$$

(A) exists, by (.01). However, for i odd, $| \text{(B)} | \geqq \left\{ \sum_{k^2=1}^{M_i} - 2 \sum_{k^2=1}^{M_{i-1}} - \sum_{k^2=1,\ \nleqq M_i}^{\infty} \right\} | a_{m_i(k)} - a_{(k)} | \geqq 3\,\delta$, and for i even,

$$| \text{(B)} | \leqq \left\{ \sum_{k^2=1}^{M_{i-1}} + \sum_{k^2=1,\ \nleqq M_i}^{\infty} \right\} | a_{m_i(k)} - a_{(k)} | \leqq 2\,\delta .$$

Hence, by (10.2) and (10.1), σ_m is not c.

11. *N.B* $\rightarrow$ *C is* (d_5). Proof similar to that of 10.

12. *N.RCRN* $\rightarrow$ *CN is* ($\bar{d}_1$). See proof of 7.

13. *N.RCURN* $\rightarrow$ *CN is* (d_2). See proof of 8.

14. *N.RC* $\rightarrow$ *CN is* (d_3). See proof of 9.

15. *N. BURCRN* $\rightarrow$ *CN is* (d_4). See 10, 12, and (.18).

16. *N.B* $\rightarrow$ *CN is* (d_5). See 11, 12, and (.19).

17. *N.RCRN* $\rightarrow$ *URC is* (e_1). (b_1) is assumed, by 3. Now $\lim_{m^1\to\infty} a_{mk}$ exists for $m^1 > D_k$ $(k = 1, 2, \cdots)$. For, with arbitrary p, $s_k \equiv \{1\ (k = p); 0$, otherwise$\}$ is rcrn, and $\sigma_m = a_{mp}$.

By denial of (e_1), there exist sequences $\delta_i > 0$, k_i, and m^1_i satisfying

$$m^1_i > m^1_{i-1}, B, D_{k_j} \qquad (j < i), \tag{17.1}$$

such that, with $(m^1 = m^1_i)$,

$$\operatorname*{osc}_{m^2=\infty} a_{(m)k_i} = \delta_i , \tag{17.2}$$

where, by (17.1),

$$\operatorname*{osc}_{m^2=\infty} a_{(m)k_j} = 0 \qquad (j < i), \tag{17.3}$$

and, by (17.1) and (b_1),

$$|a_{(m)k_j}| < A \text{ for } m^2 > B \qquad (j = 1, 2, \cdots). \tag{17.4}$$

Define $d_1 \equiv 1$; $d_i \equiv (\min_{\nu<i} d_\nu \delta_\nu)/A \cdot 2^{i+1}$ $(i > 1)$. Thus $\lim_{i\to\infty} d_i = 0$, whence $s_k \equiv \{d_i\ (k = k_i); 0$, otherwise$\}$ is rcrn. But

$$\sigma_{(m)} = \sum_{j=1}^{i-1} a_{(m)k_j} d_j \tag{A}$$

$$+ a_{(m)k_i} d_i \tag{B}$$

$$+ \sum_{j=i+1}^{\infty} a_{(m)k_j} d_j , \tag{C}$$

and $\operatorname*{osc}_{m^2=\infty}$ (A) $= 0$, by (17.3); $\operatorname*{osc}_{m^2=\infty}$ (B) $= d_i\delta_i$, by (17.2); $\operatorname*{osc}_{m^2=\infty}$ (C) $\leqq 2A \sum_{j=i+1}^{\infty} d_j \leqq d_i\delta_i/2^i$, by (17.4). Thus $\operatorname*{osc}_{m^2=\infty} \sigma_{(m)} \geqq d_i\delta_i/2$, so that by (17.1) σ_m is not urc.

18. *N.RCURN* $\rightarrow$ *URC is* (e^*_2). (b_1) is assumed, by 3. Now $\lim_{m^1\to\infty} \sum_{k^4=1}^{\infty} a_{mk}$ exists for $m^1 > E_{k^3}$ $(k^3 = 1, 2, \cdots)$. For, with arbitrary p^3, $s_k \equiv \{1\ (k^3 = p^3;$ $k^4 = 1, 2, \cdots); 0$, otherwise$\}$ is rcurn, and, with $(k^3 = p^3)$, $\sigma_m = \sum_{k^4=1}^{\infty} a_{m(k)}$.

By denial of (e^*_2), there exist π and sequences $\delta_i > 0$, k^{32}_i, and m^1_i satisfying, with $\kappa_i \equiv \pi$ $(i = 1, 2, \cdots)$,

$$m^1_i > m^1_{i-1}, B, E_{k^3_j} \qquad (j < i), \tag{18.1}$$

such that, with $(m^1 = m^1_i)$, and $[k^3 = k^3_i]$, $\operatorname*{osc}_{m^2=\infty} \sum_{k^4=1}^{\infty} a_{(m)[k]} = \delta_i$, where, by (18.1),

with $\{k^3 = k_j^3\}$, $\operatorname{osc}_{m^2=\infty} \sum_{k^4=1}^{\infty} a_{(m)\{k\}} = 0$ for $j < i$, and, by (18.1) and (b_1), $\left| \sum_{k^4=1}^{\infty} a_{(m)\{k\}} \right| < A$ for $m^2 > B$ ($j = 1, 2, \cdots$). Define $\{d_i\}$ as in the proof of 17. Then $s_k \equiv \{d_i\ (k^3 = k_i^3;\ k^4 = 1, 2, \cdots);\ 0,$ otherwise$\}$ is rcurn. But, for $m^2 > B$, $\sigma_{(m)} = \sum_{j=1}^{i-1} \sum_{k^4=1}^{\infty} a_{(m)\{k\}} d_j + \sum_{k^4=1}^{\infty} a_{(m)[k]} d_i + \sum_{j=i+1}^{\infty} \sum_{k^4=1}^{\infty} a_{(m)\{k\}} d_j$, and the conclusion follows as in the proof of 17.

19. *N.RCN → URC is* (e_2). (b_1) is assumed, by 3; and (18.1), by 18. The proof resembles that of 18.

20. *N.RC → URC is* (e_3). See proof of 9.

21. *N.BURCRN → URC is* (e_4^*). (b_1) and (e_1) are assumed, by 3 and 17, respectively. By denial of (e_4^*), there exist π and sequences m_i^1 satisfying

$$m_i^1 > m_{i-1}^1,\ B,\ D, \tag{21.1}$$

such that, with $(m^1 = m_i^1)$ and $(\kappa = \pi)$, $\overline{\lim}_{m^2 \to \infty} \sum_{k^4=1}^{\infty} | a_{(m)(k)} - a_{m_i^1(k)} | > 0$.

Hence, there exist sequences $\delta_i > 0$, and $m^2_{\nu_{ij}}$, M_{ij}, satisfying

$$B < m^2_{\nu_{11}} < m^2_{\nu_{12}} < m^2_{\nu_{21}} < m^2_{\nu_{13}} < m^2_{\nu_{22}} < m^2_{\nu_{31}} < \cdots ; \tag{21.2}$$

$$M_{11} < M_{12} < M_{21} < M_{13} < M_{22} < M_{31} < \cdots , \tag{21.3}$$

such that, with $[m^1 = m_i^1;\ m^2 = m^2_{\nu_{ij}}]$, $\sum_{k^4=1}^{M_{ij}} | a_{[m](k)} - a_{m_i^1(k)} | \geqq 6\delta_i$, where, by (21.1), (21.2), (b_1), and (.03), $\sum_{k^4=1, \leqq M_{ij}}^{\infty} | a_{[m](k)} - a_{m_i^1(k)} | \leqq \delta_i$, and, by ($e_1$), $\sum_{k^4=1}^{N_{ij}} | a_{[m](k)} - a_{m_i^1(k)} | \leqq \delta_i$, N_{ij} being the first number preceding M_{ij} in (21.3).

Now the sequence $s_k \equiv \{\operatorname{sgn}(a_{[m]k} - a_{m_i^1k})\ (\kappa = \pi;\ k^4 \leqq M_{ij}, \nleqq N_{ij}$, for j odd); 0, otherwise$\}$ is burcrn. But

$$\sigma_{[m]} = \sum_{k^4=1}^{\infty} a_{m_i^1(k)} s_{(k)} \tag{A}$$

$$+ \sum_{k^4=1}^{\infty} (a_{[m](k)} - a_{m_i^1(k)}) s_{(k)} . \tag{B}$$

(A) exists, by (21.1) and (.03). Yet for j odd, $| (B) | \geqq \left\{ \sum_{k^4=1}^{M_{ij}} - 2 \sum_{k^4=1}^{N_{ij}} - \sum_{k^4=1, \leqq M_{ij}}^{\infty} \right\} | a_{[m](k)} - a_{m_i^1(k)} | \geqq 3\,\delta_i$, while for j even,

$$| (B) | \leqq \left\{ \sum_{k^4=1}^{N_{ij}} + \sum_{k^4=1, \leqq M_{ij}}^{\infty} \right\} | a_{[m](k)} - a_{m_i^1(k)} | \leqq 2\,\delta_i,$$

whence, by (21.2), $\lim_{m^2 \to \infty} \sigma_m$ fails to exist, so that, by (21.1), σ_m is not urc.

22. *N.BCN → URC is* (e_4). (b_1) is assumed, by 3; (e_1), by 17; and (e_4^*), by 21. By denial of (e_4), there exist sequences κ_i and m_i^1 satisfying

$$m_i^1 > m_{i-1}^1, B, D, G_{\kappa_\nu} \qquad (\nu < i), \tag{22.1}$$

such that, with $(m^1 = m_i^1)$, and $(\kappa = \kappa_i)$, $\overline{\lim}_{m^2 \to \infty} \sum_{k^4=1}^{\infty} | a_{(m)(k)} - a_{m_i^1(k)} | > 0$, where, by (22.1), with $[\kappa = \kappa_\nu]$,

$$\lim_{m^2 \to \infty} \sum_{k^4=1}^{\infty} | a_{(m)[k]} - a_{m_i^1[k]} | = 0 \qquad (\nu < i). \tag{22.2}$$

Hence, there exist sequences $\delta_i > 0$, and $m_{\nu_{ij}}^2$, M_{ij} satisfying the inequalities (21.2) and (21.3), such that, with $[m^1 = m_i^1; m^2 = m_{\nu_{ij}}^2]$,

$$\sum_{k^4=1}^{M_{ij}} | a_{[m](k)} - a_{m_i(k)} | \geqq 6\, \delta_i ,$$

where, by (22.1), (21.2), (b_1), (e_1), and (.03), $\sum_{k^4=1,\, \leqq\!\!\!\!/ M_{ij}}^{\infty} | a_{[m](k)} - a_{m_i^1(k)} | \leqq \delta_i$, and, by (e_1), $\sum_{k^4=1}^{N_{ij}} | a_{[m](k)} - a_{m_i^1(k)} | \leqq \delta_i$, and, by (22.1), (b_1), and (.03),

$$\sum_{k^4=1}^{\infty} | a_{(m)[k]} - a_{m_i^1[k]} | < 2\, A \tag{22.3}$$

for $m^2 > B$, N_{ij} being defined as in the proof of 21.

Define $\{d_i\}$ as in the proof of 17. Then $s_k \equiv \{d_i \operatorname{sgn} (a_{[m]k} - a_{m_i^1k})$ $(\kappa = \kappa_i;$ $k^4 \leqq M_{ij}, \leqq\!\!\!\!/ N_{ij}$, for j odd); 0, otherwise$\}$ is bcn. Now

$$\sigma_{[m]} = \sum_{\nu=1}^{\infty} \sum_{k^4=1}^{\infty} a_{m_i^1[k]}\, s_{[k]} \tag{A}$$

$$+ \sum_{\nu=1}^{i-1} \sum_{k^4=1}^{\infty} (a_{[m][k]} - a_{m_i^1[k]})\, s_{[k]} \tag{B}$$

$$+ \sum_{k^4=1}^{\infty} (a_{[m](k)} - a_{m_i^1(k)})\, s_{(k)} \tag{C}$$

$$+ \sum_{\nu=i+1}^{\infty} \sum_{k^4=1}^{\infty} (a_{[m][k]} - a_{m_i^1[k]})\, s_{[k]} . \tag{D}$$

(A) exists, by (22.1), (21.2), and (.03); $\operatorname*{osc}_{j=\infty}$ (B) $= 0$, by (22.2) and (21.2); $\operatorname*{osc}_{j=\infty}$ (D) $\leqq 4\, A \sum_{\nu=i+1}^{\infty} d_\nu \leqq d_i \delta_i / 2^{i-1}$, by (22.3) and (21.2). But for j odd, $|\,(C)\,| \geqq \left(\sum_{k^4=1}^{M_{ij}} - 2 \sum_{k^4=1}^{N_{ij}} - \sum_{k^4=1,\, \leqq\!\!\!\!/ M_{ij}}^{\infty} \right) | a_{[m](k)} - a_{m_i^1(k)} |\, d_i \geqq 3 d_i \delta_i$, and for j even, $|\,(C)\,| \leqq \left(\sum_{k^4=1}^{N_{ij}} + \sum_{k^4=1,\, \leqq\!\!\!\!/ M_{ij}}^{\infty} \right) | a_{[m](k)} - a_{m_i^1(k)} |\, d_i \leqq 2 d_i \delta_i$, so that $\operatorname*{osc}_{j=\infty}$ (C) $\geqq$

$d_i \delta_i$, whence, by (21.2), $\operatorname*{osc}_{m^2=\infty} \sigma_{(m)} \geqq d_i\, \delta_i/2$ for $i > 1$. Thus, by (22.1), σ_m is not urc.

23. *N.B → URC is* (e_5). Proof similar to that of 21.

24. *N.RCRN → URCRN is* ($\bar{e}_1$). (b_1) and (e_1) are assumed, by 3 and 17, respectively. Now $a_{m^1k} = 0$ for $m^1 > \bar{D}_k$ ($k = 1, 2, \cdots$). (See the proof of 17.) By denial of ($\bar{e}_1$), there exist sequences $\delta_i > 0$, k_i, and m_i^1 satisfying

$$m_i^1 > m_{i-1}^1\,,\ B,\ D,\ \bar{D}_{k_j} \qquad (j < i)\,, \tag{24.1}$$

such that

$$|\, a_{m_i^1 k_i} \,| = \delta_i\,, \tag{24.2}$$

where, by (24.1),

$$a_{m_i k_j} = 0 \qquad (j < i)\,, \tag{24.3}$$

and, by (24.1) and (b_1), with ($m^1 = m_i^1$),

$$|\, a_{(m)k_j} \,| < A \text{ for } m^2 > B \qquad (j = 1, 2, \cdots)\,. \tag{24.4}$$

Define $\{d_i\}$ as in the proof of 17. Then $s_k \equiv \{d_i \operatorname{sgn} a_{m_i^1 k_i}$ ($k = k_i$); 0, otherwise$\}$ is rcrn. But σ_m is not urcrn, as can be shown by an argument analogous to that used in the conclusion to the proof of 17.

25. *N.RCURN → URCRN is* ($\bar{e}_2^*$). Proof analogous to the proof of 18 in the same way that the proof of 24 is analogous to that of 17.

26. *N.RCN → URCRN is* ($\bar{e}_2$). Proof similar to that of 19. (See remark under 25.)

27. *N.RC → URCRN is* ($\bar{e}_3$). See proof of 9.

28. *N.BURCRN → URCRN is* ($\bar{e}_4^*$). See 21, 24, and (.34).

29. *N.BCN → URCRN is* ($\bar{e}_4$). See 22, 24, and (.35).

30. *N.B → URCRN is* ($\bar{e}_5$). See 23, 24, and (.36).

31. *N.RCRN → RC is* (f_1). See proof of 7.

32. *N.RCURN → RC is* (f_2). See proof of 8.

33. *N.RC → RC is* (f_3). See proof of 9.

34. *N.BURCRN → RC is* (f_4). (c_1) and (f_1) are assumed, by 5 and 31, respectively. By denial of (f_4), there exist $\delta > 0$, r^1, π, and sequences m_i^2 ($> m_{i-1}^2$), $M_i (> M_{i-1})$, such that, with ($m^1 = r^1$; $m^2 = m_i^2$) and ($\kappa = \pi$), $\sum_{k^4=1}^{M_i} |\, a_{(m)(k)} - a_{r^1(k)} \,| \geqq 6\delta$, where, by ($c_1$), ($f_1$), and (.06), $\sum_{k^4=1, \leqq M_i}^{\infty} |\, a_{(m)(k)} - a_{r^1(k)} \,| \leqq \delta$, and, by ($f_1$), $\sum_{k^4=1}^{M_{i-1}} |\, a_{(m)(k)} - a_{r^1(k)} \,| \leqq \delta$. The remainder of the proof is similar to that of 10.

35. *N.B → RC is* (f_5). The proof is similar to that of 11.

36. *N.RCRN → RCRN is* ($\bar{f}_1$). See proof of 7.

37. *N.RCRN → RCRN is* ($\bar{f}_2$). See proof of 8.

38. *N.RC → RCRN is* ($\bar{f}_3$). See proof of 9.

39. *N.BURCRN* $\rightarrow$ *RCRN is* $(\hat{f}_4)$. See 34, 36, and (.51).
40. *N.B* $\rightarrow$ *RCRN is* $(\hat{f}_5)$. See 28, 35, and (.52).

§6. Sufficiency proofs

Proofs or indications of proofs are given only when this seems necessary.

1. *S.B* $\rightarrow$ *e is* (a_1).

2. *S.UB* $\rightarrow$ *e are* (a_1) *and* (a_2). Suppose s_k bounded for $k > Q$. Let r be arbitrary, and choose $R > \max\{Q, C_\kappa(r)\}$ $(\kappa = 1, 2, \cdots, Q)$. Then $|\sigma_r| = \left|\left\{\sum_{k=1}^{R} - \sum_{k=Q+1}^{R} + \sum_{k=Q+1}^{\infty}\right\} a_{rk}s_k\right| \leqq \left\{\sum_{k=1}^{R} + \sum_{k=Q+1}^{\infty}\right\} |a_{rk}s_k|$.

3. *S.B* $\rightarrow$ *UB is* (b_1).

4. *S.UB* $\rightarrow$ *UB are* (b_1) *and* (b_2). Suppose s_k bounded for $k > Q$. Let $R > \max\{Q, C_\kappa\}$ $(\kappa = 1, 2, \cdots, Q)$. For $m > B, R$,

$$|\sigma_m| = \left|\left\{\sum_{k=1}^{R} - \sum_{k=Q+1}^{R} + \sum_{k=Q+1}^{\infty}\right\} a_{mk}s_k\right| \leqq \left\{\sum_{k=1}^{R} + \sum_{k=Q+1}^{\infty}\right\} |a_{mk}s_k|.$$

5. *S.B* $\rightarrow$ *B is* (c_1).

6. *S.UB* $\rightarrow$ *B are* (c_1) *and* (c_2). Proof similar to that of 4.

7. *S.RCRN* $\rightarrow$ *C are* (b_1) *and* (d_1). Let R be arbitrary. By (.001), for $m > B$,

$$\sigma_m = \left\{\sum \sum_{k^1=1}^{R} \sum_{k^2=R+1}^{\infty} + \sum_{k=R+1}^{\infty}\right\} a_{mk}s_k + \sum_{k=1}^{R} a_{mk}s_k.$$

As $m, R \rightarrow \infty$, the first expression $\rightarrow 0$, by (b_1) and (.003); as $m \rightarrow \infty$, the second tends to $\sum_{k=1}^{R} a_k s_k$, and as $R \rightarrow \infty$ this expression converges, by (.01). Hence

$$\sigma = \sum_{k=1}^{\infty} a_k s_k. \tag{7.1}$$

8. *S.RCN* $\rightarrow$ *C are* (b_1), (d_1), *and* (d_2). Let R be arbitrary. For $m > B$,

$$\begin{aligned}
\sigma_m &= \left\{\sum \sum_{k^1=1}^{R} \sum_{k^2=R+1}^{\infty} + \sum_{k=1}^{R} + \sum_{k=R+1}^{\infty}\right\} a_{mk}s_k \\
&= \sum \sum_{k^1=1}^{R} \sum_{k^2=R+1}^{\infty} a_{mk}(s_k - s_{k^1}) + \sum_{k=R+1}^{\infty} a_{mk}s_k && \text{(A)} \\
&\quad + \sum \sum_{k^1=1}^{R} s_{k^1} \sum_{k^2=R+1}^{\infty} a_{mk} + \sum_{k=1}^{R} a_{mk}s_k. && \text{(B)}
\end{aligned}$$

As $m, R \rightarrow \infty$, (A) $\rightarrow 0$, by (b_1). Now by (.002), (B) $= \sum \sum_{k^1=1}^{R} s_{k^1} \left\{\sum_{k^2=1}^{\infty} - \sum_{r=1}^{t-1} (-1)^{r+1} \sum \sum_{k^{11}=1}^{R} \sum_{k^{21}=1}^{\infty} - (-1)^{t+1} \sum_{k^2=1}^{R}\right\} a_{mk} + \sum_{k=1}^{R} a_{mk}s_k$. Thus, as $m \rightarrow \infty$, (B) tends to $\sum \sum_{k^1=1}^{R} s_{k^1} \left\{L_{k^1} - \sum_{r=1}^{t-1} (-1)^{r+1} \sum \sum_{k^{11}=1}^{R} L_{k^1k^{11}} - (-1)^{t+1} \sum_{k^2=1}^{R} a_k\right\} +$

$\sum_{k=1}^{R} a_k s_k$, and as $R \to \infty$ this expression converges, by (.01) and (.02). Hence

$$\sigma = \sum \sum_{k^1=1}^{\infty} s_{k^1} \left\{ L_{k^1} - \sum_{r=1}^{t-1} (-1)^{r+1} \sum \sum_{k^{21}=1}^{\infty} L_{k^1 k^{21}} - (-1)^{t+1} \sum_{k^t=1}^{\infty} a_k \right\} + \sum_{k=1}^{\infty} a_k s_k . \tag{8.1}$$

9. *S.RC* $\to$ *C are* (b_1), (d_1), (d_2), *and* (d_3). The sequence $t_k \equiv (s_k - s)$ is rcn. Hence, by (8.1) and (d_3),

$$\sigma = s \cdot L + \sum \sum_{k^1=1}^{\infty} (s_{k^1} - s) \left\{ L_{k^1} - \sum_{r=1}^{t-1} (-1)^{r+1} \sum \sum_{k^{21}=1}^{\infty} L_{k^1 k^{21}} - (-1)^{t+1} \sum_{k^t=1}^{\infty} a_k \right\} + \sum_{k=1}^{\infty} a_k (s_k - s) . \tag{9.1}$$

10. *S.BCN* $\to$ *C are* (b_1) *and* (d_4). For $m > B$, $\sigma_m = \sum_{k=1}^{\infty} a_k s_k + \sum_{k=1}^{\infty} (a_{mk} - a_k) s_k$. These two sums exist, by (.13), (b_1), and (.01). Let R be arbitrary. Now the second sum is $\left\{ \sum_{\nu=1}^{n-1} (-1)^{\nu+1} \sum \sum_{k^1=1}^{R} \sum_{k^2=1}^{\infty} + (-1)^{n+1} \sum_{k=1}^{R} \right\} (a_{mk} - a_k) s_k + \sum_{k=R+1}^{\infty} (a_{mk} - a_k) s_k$. As $m \to \infty$, the first part of this last expression $\to 0$, by (d_4) and (.13). As $m, R \to \infty$, the second part $\to 0$, by (b_1) and (.01). Hence

$$\sigma = \sum_{k=1}^{\infty} a_k s_k . \tag{10.1}$$

11. *S.BC* $\to$ *C are* (b_1), (d_3), *and* (d_4). The sequence $t_k \equiv (s_k - s)$ is bcn. Hence, by (10.1),

$$\sigma = s \cdot L + \sum_{k=1}^{\infty} a_k (s_k - s) . \tag{11.1}$$

12. *S.B* $\to$ *C are* (b_1) *and* (d_5). For $m > B$, $\sigma_m = \sum_{k=1}^{\infty} a_k s_k + \sum_{k=1}^{\infty} (a_{mk} - a_k) s_k$, the expression existing, by (b_1), (.15), (.13), and (.01). Hence

$$\sigma = \sum_{k=1}^{\infty} a_k s_k . \tag{12.1}$$

13. *S.CN* $\to$ *C are* (b_1), (b_2), *and* (d_1). Suppose s_k bcn for $k > Q$. Let $R >$ max $\{Q, C_\kappa\}$ ($\kappa = 1, 2, \cdots, Q$). For $m > B, R$, $\sigma_m = \left\{ \sum_{k=1}^{R} - \sum_{k=Q+1}^{R} \right\} a_{mk} s_k + \sum_{k=Q+1}^{\infty} a_{mk} s_k$. As $m \to \infty$, the first expression on the right tends to $\left\{ \sum_{k=1}^{R} - \sum_{k=Q+1}^{R} \right\} a_k s_k$. By (.12), (.58), and (10.1), the second tends to $\sum_{k=Q+1}^{\infty} a_k s_k$.

Hence, by (b_2) and the nature of R,

$$\sigma = \sum_{k=1}^{\infty} a_k s_k \,. \tag{13.1}$$

14. *S.C → C are* (b_1), (b_2), (d_1), *and* (d_3). The sequence $t_k \equiv (s_k - s)$ is cn. Hence, by (13.1),

$$\sigma = s \cdot L + \sum_{k=1}^{\infty} a_k (s_k - s) \,. \tag{14.1}$$

15. *S.UB → C are* (b_1), (b_2), *and* (d_5). Suppose s_k bounded for $k > Q$. Let R be chosen as in the proof of 13, and let σ_m be decomposed in the same manner. Application of (.15) and (.13) to the first component, and of (.58) and (12.1) to the second, yield

$$\sigma = \sum_{k=1}^{\infty} a_k s_k \,. \tag{15.1}$$

16. *S.RCRN → BC are* (c_1) *and* (d_1). See 5, (.09), and 7.
17. *S.RCN → BC are* (c_1), (d_1), *and* (d_2). See 5, (.09), and 8.
18. *S.RC → BC are* (c_1), (d_1), (d_2), *and* (d_3). See 5, (.09), and 9.
19. *S.BCN → BC are* (c_1) *and* (d_4). See 5, (.09), and 10.
20. *S.BC → BC are* (c_1), (d_3), *and* (d_4). See 5, (.09), and 11.
21. *S.B → BC are* (c_1) *and* (d_5). See 5, (.09), and 12.
22. *S.CN → BC are* (c_1), (c_2), *and* (d_1). See 6, (.09), (.11), and 13.
23. *S.C → BC are* (c_1), (c_2), (d_1), *and* (d_3). See 6, (.09), (.11), and 14.
24. *S.UB → BC are* (c_1), (c_2), *and* (d_5). See 6, (.09), (.11), and 15.
25. *S.RCRN → CN are* (b_1) *and* ($\bar{d}_1$). See (.17) and (7.1).
26. *S.RCN → CN are* (b_1), ($\bar{d}_1$), *and* ($\bar{d}_2$). See (.17) and (8.1).
27. *S.RC → CN are* (b_1), ($\bar{d}_1$), ($\bar{d}_2$), *and* ($\bar{d}_3$). See (.17) and (9.1).
28. *S.BCN → CN are* (b_1) *and* ($\bar{d}_4$). See (.17), (.20), and (10.1).
29. *S.BC → CN are* (b_1), ($\bar{d}_3$), *and* ($\bar{d}_4$). See (.17), (.20), and (11.1).
30. *S.B → CN is* ($\bar{d}_5$).
31. *S.CN → CN are* (b_1), (b_2), *and* ($\bar{d}_1$). See (.17) and (13.1).
32. *S.C → CN are* (b_1), (b_2), ($\bar{d}_1$), *and* ($\bar{d}_3$). See (.17) and (14.1).
33. *S.UB → CN are* (b_2), *and* ($\bar{d}_5$). See (.17), (.76), (.23), (.20), and (15.1).
34. *S.RCRN → BCN are* (c_1) *and* ($\bar{d}_1$). See 5, (.09), and 25.
35. *S.RCN → BCN are* (c_1), ($\bar{d}_1$), *and* ($\bar{d}_2$). See 5, (.09), and 26.
36. *S.RC → BCN are* (c_1), ($\bar{d}_1$), ($\bar{d}_2$), *and* ($\bar{d}_3$). See 5, (.09), and 27.
37. *S.BCN → BCN are* (c_1) *and* ($\bar{d}_4$). See 5, (.09), and 28.
38. *S.BC → BCN are* (c_1), ($\bar{d}_3$), *and* ($\bar{d}_4$). See 5, (.09), and 29.
39. *S.B → BCN are* (c_1) *and* ($\bar{d}_5$). See 5 and 30.
40. *S.CN → BCN are* (c_1), (c_2), *and* ($\bar{d}_1$). See 6, (.09), (.11), and 31.
41. *S.C → BCN are* (c_1), (c_2), ($\bar{d}_1$), *and* ($\bar{d}_3$). See 6, (.09), (.11), and 32.
42. *S.UB → BCN are* (c_1), (c_2), *and* ($\bar{d}_5$). See 6, (.11), and 33.

43. *S.RCRN → URC are* (b_1), (d_1), *and* (e_1). The conditions are S.rcrn → c, by 7. Let $r^1 > B, D$. For $m^2 > B$, with $(m^1 = r^1)$, $\sigma_{(m)} = \left\{\sum \sum_{k^3=1}^{R} \sum_{k^4=R+1}^{\infty} + \sum_{k=R+1}^{\infty}\right\} a_{(m)k}\, s_k + \sum_{k=1}^{R} a_{(m)k}\, s_k$. As $m^2, R \to \infty$, the first part → 0, by (b_1); as $m^2 \to \infty$, the second tends to $\sum_{k=1}^{R} a_{r^1k}\, s_k$, and as $R \to \infty$, this expression converges, by (.03). Hence

$$\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k}\, s_k\,. \tag{43.1}$$

44. *S.RCURN → URC are* (b_1), (d_1), (d_2), (e_1), *and* (e_2^*). The conditions are S.rcurn → c, by 8. Suppose s_k rcrn for $k > Q$. Let $r^1 > B, D, E_\kappa$ ($\kappa = 1, 2, \cdots, Q$). For $m^2 > B$, with $(m^1 = r^1)$, $\sigma_{(m)} = \sum_{k=Q+1}^{\infty} a_{(m)k} s_k + \sum_{\nu=1}^{n-1} (-1)^{\nu+1} \sum \sum_{k^3=1}^{Q} \sum_{k^4=1}^{\infty} a_{(m)k} s_k + (-1)^{n+1} \sum_{k=1}^{Q} a_{(m)k} s_k$. By (.58), 43, and ($e_1$), as $m^2 \to \infty$, the first and third parts converge. With $(k^3 = p^3 \leqq Q)$, the summand next to the last in the second part becomes, for arbitrary R,

$$\sum_{k^4=1}^{\infty} a_{(m)(k)}\, s_{(k)} = \left\{\sum \sum_{k^{41}=1}^{R} \sum_{k^{42}=R+1}^{\infty} + \sum_{k^4=1}^{R} + \sum_{k^4=R+1}^{\infty}\right\} a_{(m)(k)}\, s_{(k)}$$

$$= \sum \sum_{k^{41}=1}^{R} \sum_{k^{42}=R+1}^{\infty} a_{(m)(k)}(s_{(k)} - s_{p^3k^{41}}) + \sum_{k^4=R+1}^{\infty} a_{(m)(k)}(s_{(k)} - s_{p^3}) \tag{A}$$

$$+ \sum \sum_{k^{41}=1}^{R} s_{p^3k^{41}} \sum_{k^{42}=R+1}^{\infty} a_{(m)(k)} + \sum_{k^4=1}^{R} a_{(m)(k)}\, s_{(k)} + s_{p^3} \sum_{k^4=R+1}^{\infty} a_{(m)(k)}\,. \tag{B}$$

As $m^2, R \to \infty$, (A) → 0, by (b_1). Now (B) $= \sum \sum_{k^{41}=1}^{R} s_{p^3k^{41}} \left\{ \sum_{k^{42}=1}^{\infty} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{421}=1}^{R} \sum_{k^{422}=1}^{\infty} - (-1)^{t+1} \sum_{k^{42}=1}^{R} \right\} a_{(m)(k)} + \sum_{k^4=1}^{R} a_{(m)(k)} s_{(k)} + s_{p^3} \left\{ \sum_{k^4=1}^{\infty} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=1}^{R} \sum_{k^{42}=1}^{\infty} - (-1)^{t+1} \sum_{k^4=1}^{R} \right\} a_{(m)(k)}$. As $m^2 \to \infty$, (B) tends to $\sum \sum_{k^{41}=1}^{R} s_{p^3k^{41}} \left\{ L_{r^1p^3k^{41}} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{421}=1}^{R} L_{r^1p^3k^{41}k^{421}} - (-1)^{t+1} \sum_{k^{42}=1}^{R} a_{r^1(k)} \right\} + \sum_{k^4=1}^{R} a_{r^1(k)} s_{(k)} + s_{p^3} \left\{ L_{r^1p^3} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=1}^{R} L_{r^1p^3k^{41}} - (-1)^{t+1} \sum_{k^4=1}^{R} a_{r^1(k)} \right\}$. As $R \to \infty$, this expression converges, by (.03) and (.04). Hence, with obvious reductions,

$$\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k} s_k + \sum_{\nu=1}^{n-1} (-1)^{\nu+1} \sum \sum_{k^3=1}^{Q} \left\{ \sum \sum_{k^{41}=1}^{\infty} s_{k^3k^{41}} \left[L_{r^1k^3k^{4:}} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \right.\right.$$

$$\sum \sum_{k^{421}=1}^{\infty} L_{r^1k^3k^{41}k^{421}} - (-1)^{t+1} \sum_{k^{42}=1}^{\infty} a_{r^1k}\Bigg] + s_{k^3}\Bigg[L_{r^1k^3} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \tag{44.1}$$
$$\sum \sum_{k^{41}=1}^{\infty} L_{r^1k^3k^{41}} - (-1)^{t+1} \sum_{k^4=1}^{\infty} a_{r^1k}\Bigg]\Bigg\}.$$

45. *S.RCN* $\rightarrow$ *URC are* (b_1), (d_1), (d_2), (e_1), *and* (e_2). The conditions are S.rcn $\rightarrow$ c, by 8. Let $r^1 > B, D, E$. For $m^2 > B$, and arbitrary R, with $(m^1 = r^1)$, $\sigma_{(m)} = \sum \sum_{k^3=1}^{R} \sum_{k^4=R+1}^{\infty} a_{(m)k}(s_k - s_{k^3}) + \sum_{k=R+1}^{\infty} a_{(m)k}\, s_k + \sum \sum_{k^3=1}^{R} s_{k^3} \sum_{k^4=R+1}^{\infty} a_{(m)k} + \sum_{k=1}^{R} a_{(m)k} s_k$. In a manner similar to that used in the proof of 8, it can be shown that

$$\sigma_{r^1} = \sum \sum_{k^3=1}^{\infty} s_{k^3}\Bigg\{L_{r^1k^3} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=1}^{\infty} L_{r^1k^3k^{41}} \tag{45.1}$$
$$- (-1)^{t+1} \sum_{k^4=1}^{\infty} a_{r^1k}\Bigg\} + \sum_{k=1}^{\infty} a_{r^1k} s_k .$$

46. *S.RC* $\rightarrow$ *URC are* (b_1), (d_1), (d_2), (d_3), (e_1), (e_2), *and* (e_3). The conditions are S.rc $\rightarrow$ c, by 9. The sequence $t_k \equiv (s_k - s)$ is rcn. Hence, by (45.1), for $r^1 > B, D, E, F$,

$$\sigma_{r^1} = s \cdot L_{r^1} + \sum \sum_{k^3=1}^{\infty} (s_{k^3} - s)\Bigg\{L_{r^1k^3} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=1}^{\infty} L_{r^1k^3k^{41}} \tag{46.1}$$
$$- (-1)^{t+1} \sum_{k^4=1}^{\infty} a_{r^1k}\Bigg\} + \sum_{k=1}^{\infty} a_{r^1k}(s_k - s) .$$

47. *S.BURCRN* $\rightarrow$ *URC are* (b_1), (d_4), (e_1), *and* (e_4^*). The conditions are S.burcrn $\rightarrow$ c, by 10. Suppose s_k rcrn for $k > Q$. Let $r^1 > B, D, G_\kappa$ $(\kappa = 1, 2, \cdots, Q)$. For $m^2 > B$, with $(m^1 = r^1)$, $\sigma_{(m)} = \sum_{k=Q+1}^{\infty} a_{(m)k}\, s_k + \sum_{\nu=1}^{n-1} (-1)^{\nu+1} \sum \sum_{k^3=1}^{Q} \sum_{k^4=1}^{\infty} a_{(m)k} s_k + (-1)^{n+1} \sum_{k=1}^{Q} a_{(m)k} s_k$. By (.58), (.13), and 43, as $m^2 \rightarrow \infty$, the first and third parts converge. With $(k^3 = p^3 \leqq Q)$, the summand next to the last in the second part becomes, for arbitrary R, $\sum_{k^4=1}^{\infty} a_{(m)(k)} s_{(k)} = \sum_{k^4=1}^{\infty} a_{r^1(k)}\, s_{(k)} + \sum_{k^4=1}^{\infty} (a_{(m)(k)} - a_{r^1(k)}) s_{(k)}$, the expression existing, by (b_1) and (.03). As $m^2 \rightarrow \infty$, the last sum $\rightarrow 0$, by (e_4^*). Hence $\sigma_{r^1} = \sum_{k=Q+1}^{\infty} a_{r^1k} s_k + \sum_{\nu=1}^{n-1} (-1)^{\nu+1} \sum \sum_{k^3=1}^{Q} \sum_{k^4=1}^{\infty} a_{r^1k} s_k + (-1)^{n+1} \sum_{k=1}^{Q} a_{r^1k} s_k$; or

$$\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k} s_k . \tag{47.1}$$

48. *S.BURCN $\to$ URC are* (b_1), (d_4), (e_1), (e_2), *and* (e_4^*). The conditions are S.burcn $\to$ c, by 10. Suppose s_k rcn for $k > Q$. Let $r^1 > B, D, E, G_\kappa$ ($\kappa = 1, 2, \cdots, Q$). For $m^2 > B$, with ($m^1 = r^1$), $\sigma_{(m)} = \sum_{k=Q+1}^{\infty} a_{(m)k}\, s_k + \sum_{\nu=1}^{n-1} (-1)^{\nu+1} \sum \sum_{k^3=1}^{Q} \sum_{k^4=1}^{\infty} a_{(m)k}\, s_k + (-1)^{n+1} \sum_{k=1}^{Q} a_{(m)k}\, s_k$. By (.58), (.13), (.14), (.64), and 45, as $m^2 \to \infty$, the first and third parts converge. With ($k^3 = p^3 \leqq Q$), the summand next to the last in the second part becomes: $\sum_{k^4=1}^{\infty} a_{(m)(k)}\, s_{(k)} = \sum_{k^4=1}^{\infty} a_{r^1(k)} s_{(k)} + \sum_{k^4=1}^{\infty} (a_{(m)(k)} - a_{r^1(k)})\, s_{(k)}$, the expression existing, by (b_1) and (.03). As $m^2 \to \infty$, the last sum $\to 0$, by (e_4^*). Hence $\sigma_{r^1} = \sum \sum_{k^3=Q+1}^{\infty} s_{k^3} \Big\{ \lim_{m^2\to\infty} \sum_{k^4=Q+1}^{\infty} a_{(m)k} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=Q+1}^{\infty} \lim_{m^2\to\infty} \sum_{k^{42}=Q+1}^{\infty} a_{(m)k} - (-1)^{t+1} \sum_{k^4=Q+1}^{\infty} a_{r^1k} \Big\} + \sum_{k=1}^{\infty} a_{r^1k} s_k$, which reduces to

$$
\begin{aligned}
\sigma_{r^1} = \sum \sum_{k^3=Q+1}^{\infty} s_{k^3} \Big\{ & L_{r^1k^3} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=1}^{Q} L_{r^1k^3k^{41}} - (-1)^{t+1} \\
& \sum_{k^4=1}^{Q} a_{r^1k} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=Q+1}^{\infty} \Big[L_{r^1k^3k^{41}} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{421}=1}^{Q} L_{r^1k^3k^{41}k^{421}} \\
& - (-1)^{t+1} \sum_{k^{42}=1}^{Q} a_{r^1k} \Big] - (-1)^{t+1} \sum_{k^4=Q+1}^{\infty} a_{r^1k} \Big\} + \sum_{k=1}^{\infty} a_{r^1k}\, s_k\,.
\end{aligned} \tag{48.1}
$$

49. *S.BURC $\to$ URC are* (b_1), (d_3), (d_4), (e_1), (e_2), (e_3), *and* (e_4^*). The conditions are S.burc $\to$ c, by 11. The sequence $t_k \equiv (s_k - s)$ is burcn. Hence, by (48.1), for $r^1 > B, D, E, F, G_\kappa$ ($\kappa = 1, 2, \cdots, Q$), Q being defined as in the proof of 48,

$$
\begin{aligned}
\sigma_{r^1} = s \cdot L_{r^1} + \sum \sum_{k^3=Q+1}^{\infty} (s_{k^3} - s) \Big\{ & L_{r^1k^3} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=1}^{Q} L_{r^1k^3k^{41}} \\
& - (-1)^{t+1} \sum_{k^4=1}^{Q} a_{r^1k} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=Q+1}^{\infty} \\
& \Big[L_{r^1k^3k^{41}} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{421}=1}^{Q} L_{r^1k^3k^{41}k^{421}} - (-1)^{t+1} \sum_{k^{42}=1}^{Q} a_{r^1k} \Big] \\
& - (-1)^{t+1} \sum_{k^4=Q+1}^{\infty} a_{r^1k} \Big\} + \sum_{k=1}^{\infty} a_{r^1k} (s_k - s)\,.
\end{aligned} \tag{49.1}
$$

50. *S.BCN $\to$ URC are* (b_1), (d_4), *and* (e_4). The conditions are S.bcn $\to$ c, by 10. For $r^1 > B, H$, it can be proved in a manner similar to that used in the proof of 10, that

$$
\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k}\, s_k\,. \tag{50.1}
$$

51. *S.BC* $\rightarrow$ *URC are* (b_1), (d_3), (d_4), (e_3), *and* (e_4). The conditions are S.bc $\rightarrow$ c, by 11. The sequence $t_k \equiv (s_k - s)$ is bcn. Hence, by (50.1), for $r^1 > B, F, H$,

$$\sigma_{r^1} = s \cdot L + \sum_{k=1}^{\infty} a_{r^1k} (s_k - s) . \tag{51.1}$$

52. *S.B* $\rightarrow$ *URC are* (b_1), (d_5), *and* (e_5). The conditions are S.b $\rightarrow$ c, by 12. For $r^1 > B, I$, it can be proved in a manner similar to that used in the proof of 12 that

$$\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k} \, s_k . \tag{52.1}$$

53. *S.URCRN* $\rightarrow$ *URC are* (b_1), (b_2), (d_1), *and* (e_1). The conditions are S.urcrn $\rightarrow$ c, by 13. Suppose s_k rcn for $k > Q$. Let $R > \max \{Q, C_\kappa\}$ $(\kappa = 1, 2, \cdots, Q)$. For $r^1 > B, D, R$, it can be proved in a manner similar to that used in the proof of 13, that

$$\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k} \, s_k . \tag{53.1}$$

54. *S.URCN* $\rightarrow$ *URC are* (b_1), (b_2), (d_1), (e_1) *and* (e_2). The conditions are S.urcn $\rightarrow$ c, by 13. Suppose s_k rcn for $k > Q$. Let $R > \max \{Q, C_\kappa\}$ $(\kappa = 1, 2, \cdots, Q)$. Let $r^1 > B, D, E, R$. For $m^2 > B, R$, with $(m^1 = r^1)$, $\sigma_{(m)} = \left\{ \sum_{k=1}^{R} - \sum_{k=Q+1}^{R} \right\} a_{(m)k} \, s_k + \sum_{k=Q+1}^{\infty} a_{(m)k} \, s_k$. As $m^2 \rightarrow \infty$, the first part tends to $\left\{ \sum_{k=1}^{R} - \sum_{k=Q+1}^{R} \right\} a_{r^1k} \, s_k$. By (.58), (.12), (.14), (.64), and (45.1), the second tends to $\sum \sum_{k^\delta=Q+1}^{\infty} s_{k^\delta} \left\{ \lim_{m^2\rightarrow\infty} \sum_{k^\epsilon=Q+1}^{\infty} a_{(m)k} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{\epsilon_1}=Q+1}^{\infty} \lim_{m^2\rightarrow\infty} \sum_{k^{\epsilon_2}=Q+1}^{\infty} a_{(m)k} - (-1)^{t+1} \sum_{k^\epsilon=Q+1}^{\infty} a_{r^1k} \right\} + \sum_{k=Q+1}^{\infty} a_{r^1k} \, s_k$. Hence, as in the proof of 48,

$$\begin{aligned} \sigma_{r^1} = \sum \sum_{k^\delta=Q+1}^{\infty} s_{k^\delta} \Bigg\{ & L_{r^1k^\delta} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{\epsilon_1}=1}^{Q} L_{r^1k^\delta k^{\epsilon_1}} \\ & - (-1)^{t+1} \sum_{k^\epsilon=1}^{Q} a_{r^1k} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{\epsilon_1}=Q+1}^{\infty} \\ & \left[L_{r^1k^\delta k^{\epsilon_1}} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{\epsilon_2}=1}^{Q} L_{r^1k^\delta k^{\epsilon_1}k^{\epsilon_2}} - (-1)^{t+1} \sum_{k^{\epsilon_1}=1}^{Q} a_{r^1k} \right] \\ & - (-1)^{t+1} \sum_{k^\epsilon=Q+1}^{\infty} a_{r^1k} \Bigg\} + \sum_{k=1}^{\infty} a_{r^1k} \, s_k . \end{aligned} \tag{54.1}$$

55. *S.URC = URC are* (b_1), (b_2), (d_1), (d_3), (e_1), (e_2), *and* (e_3). The conditions are S.urc $\to$ c, by 14. The sequence $t_k \equiv (s_k - s)$ is urcn. Hence, by (54.1), for $r^1 > B, D, E, R$,

$$
\begin{aligned}
\sigma_{r^1} = s \cdot L_{r^1} + \sum \sum_{k^3=Q+1}^{\infty} (s_{k^3} - s) \Big\{ L_{r^1k^3} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=1}^{Q} L_{r^1k^3k^{41}} \\
-(-1)^{t+1} \sum_{k^4=1}^{Q} a_{r^1k} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \sum \sum_{k^{41}=Q+1}^{\infty} \Big[L_{r^1k^3k^{41}} - \sum_{\tau=1}^{t-1} (-1)^{\tau+1} \\
\sum \sum_{k^{421}=1}^{Q} L_{r^1k^3k^{41}k^{421}} - (-1)^{t+1} \sum_{k^{42}=1}^{Q} a_{r^1k} \Big] - (-1)^{t+1} \sum_{k^4=Q+1}^{\infty} a_{r^1k} \Big\} \\
+ \sum_{k=1}^{\infty} a_{r^1k} (s_k - s) .
\end{aligned}
\tag{55.1}
$$

56. *S.CN $\to$ URC are* (b_1), (b_2), (d_1), *and* (e_4). The conditions are S.cn $\to$ c, by 13. Suppose s_k bcn for $k > Q$. Let $R > \max \{Q, C_\kappa\}$ $(\kappa = 1, 2, \cdots, Q)$. For $r^1 > B, H, R$, it can be proved in a manner similar to that used in the proof of 13 that

$$
\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k} s_k . \tag{56.1}
$$

57. *S.C $\to$ URC are* (b_1), (b_2), (d_1), (d_3), (e_3), *and* (e_4). The conditions are S.c $\to$ c, by 14. The sequence $t_k \equiv (s_k - s)$ is cn. Hence, by (56.1), for $r^1 > B, H, R$, where R is defined as in the proof of 56,

$$
\sigma_{r^1} = s \cdot L + \sum_{k=1}^{\infty} a_{r^1k} (s_k - s) . \tag{57.1}
$$

58. *S.UB $\to$ URC are* (b_1), (b_2), (d_5), *and* (e_5). The conditions are S.ub $\to$ c, by 15. Suppose s_k bounded for $k > Q$. Let $R > \max \{Q, C_\kappa\}$ $(\kappa = 1, 2, \cdots, Q)$. For $r^1 > B, I, R$, it can be proved in a manner similar to that used in the proof of 15, that

$$
\sigma_{r^1} = \sum_{k=1}^{\infty} a_{r^1k} s_k . \tag{58.1}
$$

59. *S.RCRN $\to$ BURC are* (c_1), (d_1), *and* (e_1). See 5, (.09), and 43.

60. *S.RCURN $\to$ BURC are* (c_1), (d_1), (d_2), (e_1), *and* (e_2^*). See 5, (.09), and 44.

61. *S.RCN $\to$ BURC are* (c_1), (d_1), (d_2), (e_1), *and* (e_2). See 5, (.09), and 45.

62. *S.RC $\to$ BURC are* (c_1), (d_1), (d_2), (d_3), (e_1), (e_2), *and* (e_3). See 5, (.09), and 46.

63. *S.BURCRN $\to$ BURC are* (c_1), (d_4), (e_1), *and* (e_4^*). See 5, (.09), and 47.

64. *S.BURCN $\to$ BURC are* (c_1), (d_4), (e_1), (e_2), *and* (e_4^*). See 5, (.09), and 48.

65. *S.BURC $\to$ BURC are* (c_1), (d_3), (d_4), (e_1), (e_2), (e_3), *and* (e_4^*). See 5, (.09), and 49.

66. *S.BCN* → *BURC are* (c_1), (d_4), *and* (e_4). See 5, (.09), and 50.

67. *S.BC* → *BURC are* (c_1), (d_3), (d_4), (e_3), *and* (e_4). See 5, (.09), and 51.

68. *S.B* → *BURC are* (c_1), (d_5), *and* (e_5). See 5, (.09), and 52.

69. *S.CN* → *BURC are* (c_1), (c_2), (d_1), *and* (e_1). See 6, (.09), (.11), (.25), and 56.

70. *S.C* → *BURC are* (c_1), (c_2), (d_1), (d_3), (e_1), *and* (e_3). See 6, (.09), (.11), (.25), and 57.

71. *S.UB* → *BURC are* (c_1), (c_2), (d_5), *and* (e_5). See 6, (.09), (.11), and 58.

72. *S.RCRN* → *URCN and* (b_1), ($\bar{d}_1$), *and* (e_1). See (.17), 43, and 25.

73. *S.RCURN* → *URCN are* (b_1), ($\bar{d}_1$), ($\bar{d}_2$), (e_1), *and* (e_2^*). See (.17), 44, and 26.

74. *S.RCN* → *URCN are* (b_1), ($\bar{d}_1$), ($\bar{d}_2$), (e_1), *and* (e_2). See (.17), 45, and 26.

75. *S.RC* → *URCN are* (b_1), ($\bar{d}_1$), ($\bar{d}_2$), ($\bar{d}_3$), (e_1), (e_2), *and* (e_3). See (.17), 46, and 27.

76. *S.BURCRN* → *URCN are* (b_1), ($\bar{d}_4$), (e_1), *and* (e_4^*). See (.17), 47, and 28.

77. *S.BURCN* → *URCN are* (b_1), ($\bar{d}_4$), (e_1), (e_2), *and* (e_4^*). See (.17), 48, and 28.

78. *S.BURC* → *URCN are* (b_1), ($\bar{d}_3$), ($\bar{d}_4$), (e_1), (e_2), (e_3), *and* (e_4^*). See (.17), 49, and 29.

79. *S.BCN* → *URCN are* (b_1), ($\bar{d}_4$), *and* (e_4). See (.17), 50, and 28.

80. *S.BC* → *URCN are* (b_1), (d_3), ($\bar{d}_4$), (e_3), *and* (e_4). See (.17), 51, and 29.

81. *S.B* → *URCN are* ($\bar{d}_5$) *and* (e_5). See (.17), (.76), 52, and 30.

82. *S.URCRN* → *URCN are* (b_1), (b_2), ($\bar{d}_1$), *and* (e_1). See (.17), 53, and 31.

83. *S.URCN* → *URCN are* (b_1), (b_2), ($\bar{d}_1$), (e_1), *and* (e_2). See (.17), 54, and 31.

84. *S.URC* → *URCN are* (b_1), (b_2), ($\bar{d}_1$), ($\bar{d}_3$), (e_1), (e_2), *and* (e_3). See (.17), 55, and 32.

85. *S.CN* → *URCN are* (b_1), (b_2), ($\bar{d}_1$), *and* (e_4). See (.17), 56, and 31.

86. *S.C* → *URCN are* (b_1), (b_2), ($\bar{d}_1$), ($\bar{d}_3$), (e_3), *and* (e_4). See (.17), 57, and 32.

87. *S.UB* → *URCN are* (b_2), ($\bar{d}_5$), *and* (e_5). See (.17), (.76), 58, and 33.

88. *S.RCRN* → *BURCN are* (c_1), ($\bar{d}_1$), *and* (e_1). See 5, (.09), and 72.

89. *S.RCURN* → *BURCN are* (c_1), ($\bar{d}_1$), ($\bar{d}_2$), (e_1), *and* (e_2^*). See 5, (.09), and 73.

90. *S.RCN* → *BURCN are* (c_1), ($\bar{d}_1$), ($\bar{d}_2$), (e_1), *and* (e_2). See 5, (.09), and 74.

91. *S.RC* → *BURCN are* (c_1), ($\bar{d}_1$), ($\bar{d}_2$), ($\bar{d}_3$), (e_1), (e_2), *and* (e_3). See 5, (.09), and 75.

92. *S.BURCRN* → *BURCN are* (c_1), ($\bar{d}_4$), (e_1), *and* (e_4^*). See 5, (.09), and 76.

93. *S.BURCN* → *BURCN are* (c_1), ($\bar{d}_4$), (e_1), (e_2), *and* (e_4^*). See 5, (.09), and 77.

94. *S.BURC* → *BURCN are* (c_1), ($\bar{d}_3$), ($\bar{d}_4$), (e_1), (e_2), (e_3), *and* (e_4^*). See 5, (.09), and 78.

95. *S.BCN* → *BURCN are* (c_1), ($\bar{d}_4$), *and* (e_4). See 5, (.09), and 79.

96. *S.BC* → *BURCN are* (c_1), ($\bar{d}_3$), ($\bar{d}_4$), (e_3), *and* (e_4). See 5, (.09), and 80.

97. *S.B* → *BURCN are* (c_1), ($\bar{d}_5$), *and* (e_5). See 5 and 81.

98. *S.CN* → *BURCN are* (c_1), (c_2), ($\bar{d}_1$), *and* (e_1). See 6, (.09), (.11), (.25), and 85.

99. $S.C \rightarrow BURCN$ *are* (c_1), (c_2), $(\bar{d}_1)$, $(\bar{d}_3)$, (e_1), *and* (e_3). See 6, (.09), (.11), (.25), and 86.

100. $S.UB \rightarrow BURCN$ *are* (c_1), (c_2), $(\bar{d}_5)$, *and* (e_5). See 6, (.11), and 87.

101. $S.RCRN \rightarrow URCRN$ *are* (b_1), (d_1), *and* $(\bar{e}_1)$. See (.33) and (43.1).

102. $S.RCURN \rightarrow URCRN$ *are* (b_1), (d_1), (d_2), $(\bar{e}_1)$, *and* $(\bar{e}_2^*)$. See (.33) and (44.1).

103. $S.RCN \rightarrow URCRN$ *are* (b_1), (d_1), (d_2), $(\bar{e}_1)$, *and* $(\bar{e}_2)$. See (.33) and (45.1).

104. $S.RC \rightarrow URCRN$ *are* (b_1), (d_1), (d_2), (d_3), $(\bar{e}_1)$, $(\bar{e}_2)$, *and* $(\bar{e}_3)$. See (.33) and (46.1).

105. $S.BURCRN \rightarrow URCRN$ *are* (b_1), (d_4), (e_4^*), *and* $(\bar{e}_1)$. See (.33) and (47.1).

106. $S.BURCN \rightarrow URCRN$ *are* (b_1), (d_4), (e_4^*), $(\bar{e}_1)$, *and* $(\bar{e}_2)$. See (.33) and (48.1).

107. $S.BURC \rightarrow URCRN$ *are* (b_1), (d_3), (d_4), (e_4^*), $(\bar{e}_1)$, $(\bar{e}_2)$, *and* $(\bar{e}_3)$. See (.33) and (49.1).

108. $S.BCN \rightarrow URCRN$ *are* (b_1), (d_4), *and* $(\bar{e}_4)$. See (.33) and (50.1).

109. $S.BC \rightarrow URCRN$ *are* (b_1), (d_3), (d_4), $(\bar{e}_3)$, *and* $(\bar{e}_4)$. See (.33) and (51.1).

110. $S.B \rightarrow URCRN$ *are* (d_5) *and* $(\bar{e}_5)$. See (.33), (.43), (.39), (.15), (.13), (.73), (.19), (.76), and (52.1).

111. $S.URCRN \rightarrow URCRN$ *are* (b_1), (b_2), (d_1), *and* $(\bar{e}_1)$. See (.33) and (53.1).

112. $S.URCN \rightarrow URCRN$ *are* (b_1), (b_2), (d_1), $(\bar{e}_1)$, *and* $(\bar{e}_2)$. See (.33) and (54.1).

113. $S.URC \rightarrow URCRN$ *are* (b_1), (b_2), (d_1), (d_3), $(\bar{e}_1)$, $(\bar{e}_2)$, *and* $(\bar{e}_3)$. See (.33) and (55.1).

114. $S.CN \rightarrow URCRN$ *are* (b_1), (b_2), (d_1), *and* $(\bar{e}_4)$. See (.33) and (56.1).

115. $S.C \rightarrow URCRN$ *are* (b_1), (b_2), (d_1), (d_3), $(\bar{e}_3)$, *and* $(\bar{e}_4)$. See (.33) and (57.1).

116. $S.UB \rightarrow URCRN$ *are* (b_2), (d_5), *and* $(\bar{e}_5)$. See (.33), (.43), (.39), (.15), (.13), (.73), (.19), (.76), and (58.1).

117. $S.RCRN \rightarrow BURCRN$ *are* (c_1), (d_1), *and* $(\bar{e}_1)$. See 5, (.09), and 101.

118. $S.RCURN \rightarrow BURCRN$ *are* (c_1), (d_1), (d_2), $(\bar{e}_1)$, *and* $(\bar{e}_2^*)$. See 5, (.09), and 102.

119. $S.RCN \rightarrow BURCRN$ *are* (c_1), (d_1), (d_2), $(\bar{e}_1)$, *and* $(\bar{e}_2)$. See 5, (.09), and 103.

120. $S.RC \rightarrow BURCRN$ *are* (c_1), (d_1), (d_2), (d_3), $(\bar{e}_1)$, $(\bar{e}_2)$, *and* $(\bar{e}_3)$. See 5, (.09), and 104.

121. $S.BURCRN \rightarrow BURCRN$ *are* (c_1), (d_4), (e_4^*), *and* $(\bar{e}_1)$. See 5, (.09), and 105.

122. $S.BURCN \rightarrow BURCRN$ *are* (c_1), (d_4), (e_4^*), $(\bar{e}_1)$, *and* $(\bar{e}_2)$. See 5, (.09), and 106.

123. $S.BURC \rightarrow BURCRN$ *are* (c_1), (d_3), (d_4), (e_4^*), $(\bar{e}_1)$, $(\bar{e}_2)$, *and* $(\bar{e}_3)$. See 5, (.09), and 107.

124. $S.BCN \rightarrow BURCRN$ *are* (c_1), (d_4), *and* $(\bar{e}_4)$. See 5, (.09), and 108.

125. *S.BC* $\rightarrow$ *BURCRN are* (c_1), (d_3), (d_4), ($\bar{e}_3$), *and* ($\bar{e}_4$). See 5, (.09), and 109.

126. *S.B* $\rightarrow$ *BURCRN are* (c_1), (d_5), *and* ($\bar{e}_5$). See 5 and 110.

127. *S.CN* $\rightarrow$ *BURCRN are* (c_1), (c_2), (d_1), *and* ($\bar{e}_1$). See 6, (.09), (.11), (.33), (.25), (.35), and 114.

128. *S.C* $\rightarrow$ *BURCRN are* (c_1), (c_2), (d_1), (d_3), ($\bar{e}_1$), *and* ($\bar{e}_3$). See 6, (.09), (.11), (.33), (.25), (.35), and 115.

129. *S.UB* $\rightarrow$ *BURCRN are* (c_1), (c_2), (d_5), *and* ($\bar{e}_5$). See 6, (.11), and 116.

The proofs of ## 130–138 have already been constructed under the corresponding theorems for urc transformations, among ## 43–58, as examination of these proofs will show, in view of (.09), (.11), (.44), and (.49). Also, the form of the row-limits of the transform will be found there.

130. *S.RCRN* $\rightarrow$ *RC are* (c_1), (d_1), *and* (f_1).

131. *S.RCN* $\rightarrow$ *RC are* (c_1), (d_1), (d_2), (f_1), *and* (f_2).

132. *S.RC* $\rightarrow$ *RC are* (c_1), (d_1), (d_2), (d_3), (f_1), (f_2), *and* (f_3).

133. *S.BCN* $\rightarrow$ *RC are* (c_1), (d_4), *and* (f_4).

134. *S.BC* $\rightarrow$ *RC are* (c_1), (d_3), (d_4), (f_3), *and* (f_4).

135. *S.B* $\rightarrow$ *RC are* (c_1), (d_5), *and* (f_5).

136. *S.CN* $\rightarrow$ *RC are* (c_1), (c_2), (d_1), *and* (f_1).

137. *S.C* $\rightarrow$ *RC are* (c_1), (c_2), (d_1), (d_3), (f_1), *and* (f_3).

138. *S.UB* $\rightarrow$ *RC are* (c_1), (c_2), (d_5), *and* (f_5).

139. *S.RCRN* $\rightarrow$ *RCN are* (c_1), ($\bar{d}_1$), *and* (f_1). See (.17), 130, and 34.

140. *S.RCN* $\rightarrow$ *RCN are* (c_1), ($\bar{d}_1$), ($\bar{d}_2$), (f_1), *and* (f_2). See (.17), 131, and 35.

141. *S.RC* $\rightarrow$ *RCN are* (c_1), ($\bar{d}_1$), ($\bar{d}_2$), ($\bar{d}_3$), (f_1), (f_2), *and* (f_3). See (.17), 132, and 36.

142. *S.BCN* $\rightarrow$ *RCN are* (c_1), ($\bar{d}_4$), *and* (f_4). See (.17), 133, and 37.

143. *S.BC* $\rightarrow$ *RCN are* (c_1), ($\bar{d}_3$), ($\bar{d}_4$), (f_3), *and* (f_4). See (.17), 134, and 38.

144. *S.B* $\rightarrow$ *RCN are* (c_1), ($\bar{d}_5$), *and* (f_5). See (.17), 135, and 39.

145. *S.CN* $\rightarrow$ *RCN are* (c_1), (c_2), ($\bar{d}_1$), *and* (f_1). See (.17), 136, and 40.

146. *S.C* $\rightarrow$ *RCN are* (c_1), (c_2), ($\bar{d}_1$), ($\bar{d}_3$), (f_1), *and* (f_3). See (.17), 137, and 41.

147. *S.UB* $\rightarrow$ *RCN are* (c_1), (c_2), ($\bar{d}_5$), *and* (f_5). See (.17), 138, and 42.

148. *S.RCRN* $\rightarrow$ *RCURN are* (c_1), (d_1), ($\bar{e}_1$), *and* (f_1). See 130 and 117.

149. *S.RCURN* $\rightarrow$ *RCURN are* (c_1), (d_1), (d_2), ($\bar{e}_1$), ($\bar{e}_2^*$), (f_1), *and* (f_2). See 131 and 118.

150. *S.RCN* $\rightarrow$ *RCURN are* (c_1), (d_1), (d_2), ($\bar{e}_1$), ($\bar{e}_2$), (f_1), *and* (f_2). See 131 and 119.

151. *S.RC* $\rightarrow$ *RCURN are* (c_1), (d_1), (d_2), (d_3), ($\bar{e}_1$), ($\bar{e}_2$), ($\bar{e}_3$), (f_1), (f_2), *and* (f_3). See 132 and 120.

152. *S.BCN* $\rightarrow$ *RCURN are* (c_1), (d_4), ($\bar{e}_1$), *and* (f_4). See 133, (.49), (.35), and 124.

153. *S.BC* $\rightarrow$ *RCURN are* (c_1), (d_3), (d_4), ($\bar{e}_1$), ($\bar{e}_3$), (f_3), *and* (f_4). See 134, (.49), (.35), and 125.

154. *S.B* $\rightarrow$ *RCURN are* (c_1), (d_5), ($\bar{e}_1$), *and* (f_5). See 135, (.49), (.36), and 126.

155. *S.CN* → *RCURN are* (c_1), (c_2), (d_1), ($\bar{e}_1$), *and* (f_1). See 136 and 127.

156. *S.C* → *RCURN are* (c_1), (c_2), (d_1), (d_3), ($\bar{e}_1$), ($\bar{e}_3$), (f_1), *and* (f_3). See 137 and 128.

157. *S.UB* → *RCURN are* (c_1), (c_2), (d_5), ($\bar{e}_1$), *and* (f_5). See 138, (.49), (.36), and 129.

158. *S.RCRN* → *RCRN are* (c_1), (d_1), *and* ($\mathring{f}_1$). See (.50), 130, and (43.1).[13]

159. *S.RCN* → *RCRN are* (c_1), (d_1), (d_2), ($\mathring{f}_1$), *and* ($\mathring{f}_2$). See (.50), 131, and (45.1).[13]

160. *S.RC* → *RCRN are* (c_1), (d_1), (d_2), (d_3), ($\mathring{f}_1$), ($\mathring{f}_2$), *and* ($\mathring{f}_3$). See (.50), 132, and (46.1).[13]

161. *S.BCN* → *RCRN are* (c_1), (d_4), *and* ($\mathring{f}_4$). See (.50), 133, and (50.1).[13]

162. *S.BC* → *RCRN are* (c_1), (d_3), (d_4), ($\mathring{f}_3$), *and* ($\mathring{f}_4$). See (.50), 134, and (51.1).[13]

163. *S.B* → *RCRN are* (a_1), (d_5), *and* ($\mathring{f}_5$). See (.50), (.57), (.43), (.39), (.15), (.13), (.73), (.19), (.77), 135, and (52.1).[13]

164. *S.CN* → *RCRN are* (c_1), (c_2), (d_1), *and* ($\mathring{f}_1$). See (.50), 136, and (56.1).[13]

165. *S.C* → *RCRN are* (c_1), (c_2), (d_1), (d_3), ($\mathring{f}_1$), *and* ($\mathring{f}_3$). See (.50), 137, and (57.1).[13]

166. *S.UB* → *RCRN are* (a_1), (c_2), (d_5), *and* ($\mathring{f}_5$). See (.50), (.57), (.43), (.39), (.15), (.13), (.73), (.19), (.77), 138, and (58.1).[13]

§7. Conclusion

It is now possible to write down at once S. conditions for the transformation $\{s_k\} \to \{\sigma_m\}$ in each of the 256 possible cases. For example, consider burcn → cn. Now ## 25–33 of §6 state conditions S. that the transforms of various sequences $\{s_k\}$ be cn, and among them # 28 involves the most restricted class in these ## (namely bcn) including the type burcn. (See diagram.) Hence S.burcn → cn are (b_1) and ($\bar{d}_4$).

The S. conditions thus secured will be found to be always N. for the transformation concerned, under the added hypothesis of existence. Thus in the above example (b_1) is N.burcn → (cn and e) by # 3 of §5, and ($\bar{d}_4$) is N.burcn → cn by # 15 of §5. (See diagram.)

From the S. conditions thus obtained it is possible to write at once conditions S. that the transform be also existent. This can be accomplished by adding (a_1) if $\{s_k\}$ is rcrn, rcurn, rcn, rc, burcrn, burcn, burc, bcn, bc, or b, and by adding (a_1) and (a_2) if $\{s_k\}$ is urcrn, urcn, urc, cn, c, or ub, as is shown by ## 1 and 2 of §6. Also, ## 1 and 2 of §5 show that these conditions are N. for existence. Hence the set of conditions thus obtained will be both N. and S. It is to be observed that, by (.08) and (.10), the addition of (a_1) is unnecessary when (c_1) already appears, and the addition of (a_2) is superfluous when (c_2) is present. Thus N. and S. burcn → (cn and e) are (a_1), (b_1), and ($\bar{d}_4$).

When $\|a_{mk}\|$ is row-finite, (a_1) and (a_2) are automatically satisfied, so that the question of existence does not enter. Thus for row-finite (and, in particular,

[13] See remark preceding 130.

for triangular) matrices, N. and S. burcn $\rightarrow$ cn are (b_1) and $(\bar{d}_4)$, and σ_m exists for each m.

Stronger sets of S. conditions for any transformation can easily be secured, as suggested in §2. Thus conditions S.U $\rightarrow$ v, where U, v represent classes of sequences, are S.burcn $\rightarrow$ cn whenever U $\supset$ burcn and cn $\supset$ v. To find all of those conditions in §3 which are N. for a given transformation involves either applying the relations in §4 to the N. conditions already obtained, or collecting all relevant implications from the theorems in §5. Thus N.burcn $\rightarrow$ (cn and e) are (a_1), (b_1), (d_1), (d_2), (d_4), $(\bar{d}_1)$, $(\bar{d}_2)$, and $(\bar{d}_4)$.

Conditions for convergence preservation, with preservation of the limit for *null sequences*, are also available in the several cases as combinations of the corresponding separate sets of conditions. In view of (.17), these may be simplified to yield the following table, which is here presented for convenience in possible future reference.

RC	$\rightarrow$ C	(a_1), (b_1), (d_3), $(\bar{d}_1)$, and $(\bar{d}_2)$.
BURC	$\rightarrow$ C	(a_1), (b_1), (d_3), and $(\bar{d}_4)$.
BC	$\rightarrow$ C	(a_1), (b_1), (d_3), and $(\bar{d}_4)$.
URC	$\rightarrow$ C	(a_1), (a_2), (b_1), (b_2), (d_3), and $(\bar{d}_1)$.
C	$\rightarrow$ C	(a_1), (a_2), (b_1), (b_2), (d_3), and $(\bar{d}_1)$.
RC	$\rightarrow$ BC	(c_1), (d_3), $(\bar{d}_1)$, and $(\bar{d}_2)$.
BURC	$\rightarrow$ BC	(c_1), (d_3), and $(\bar{d}_4)$.
BC	$\rightarrow$ BC	(c_1), (d_3), and $(\bar{d}_4)$.
URC	$\rightarrow$ BC	(c_1), (c_2), (d_3), and $(\bar{d}_1)$.
C	$\rightarrow$ BC	(c_1), (c_2), (d_3), and $(\bar{d}_1)$.
RC	$\rightarrow$ URC	(a_1), (b_1), (d_3), $(\bar{d}_1)$, $(\bar{d}_2)$, (e_1), (e_2), and (e_3).
BURC	$\rightarrow$ URC	(a_1), (b_1), (d_3), $(\bar{d}_4)$, (e_1), (e_2), (e_3), and (e_4^*).
BC	$\rightarrow$ URC	(a_1), (b_1), (d_3), $(\bar{d}_4)$, (e_3), and (e_4).
URC	$\rightarrow$ URC	(a_1), (a_2), (b_1), (b_2), (d_3), $(\bar{d}_1)$, (e_1), (e_2), and (e_3).
C	$\rightarrow$ URC	(a_1), (a_2), (b_1), (b_2), (d_3), $(\bar{d}_1)$, (e_3), and (e_4).
RC	$\rightarrow$ BURC	(c_1), (d_3), $(\bar{d}_1)$, $(\bar{d}_2)$, (e_1), (e_2), and (e_3).
BURC	$\rightarrow$ BURC	(c_1), (d_3), $(\bar{d}_4)$, (e_1), (e_2), (e_3), and (e_4^*).
BC	$\rightarrow$ BURC	(c_1), (d_3), $(\bar{d}_4)$, (e_3), and (e_4).
URC	$\rightarrow$ BURC	(c_1), (c_2), (d_3), $(\bar{d}_1)$, (e_1), and (e_3).
C	$\rightarrow$ BURC	(c_1), (c_2), (d_3), $(\bar{d}_1)$, (e_1), and (e_3).
RC	$\rightarrow$ RC	(c_1), (d_3), $(\bar{d}_1)$, $(\bar{d}_2)$, (f_1), (f_2), and (f_3).
BURC	$\rightarrow$ RC	(c_1), (d_3), $(\bar{d}_4)$, (f_3), and (f_4).
BC	$\rightarrow$ RC	(c_1), (d_3), $(\bar{d}_4)$, (f_3), and (f_4).
URC	$\rightarrow$ RC	(c_1), (c_2), (d_3), $(\bar{d}_1)$, (f_1), and (f_3).
C	$\rightarrow$ RC	(c_1), (c_2), (d_3), $(\bar{d}_1)$, (f_1), and (f_3).

These sets are N. and S., existence being assumed. If (a_1) and (a_2) are deleted, it may be again remarked, existence is not necessarily implied (save when the transform is bounded).

Now by examination of (9.1), (11.1), and (14.1), it is seen that in each case $\sigma = s \cdot L$. If, then, $L = 1$ in the formulation of (d_3), N. and S. conditions for regularity are secured.

Brown University.

THE GROUPS DETERMINED BY THE RELATIONS

$S^l = T^m = (S^{-1}T^{-1}ST)^p = 1$

PART I

BY H. S. M. COXETER

In working out the commutator subgroups of the finite groups generated by reflections, I came across a group of order 288 having the abstract definition

$$S^3 = T^3 = (S^{-1}T^{-1}ST)^2 = 1.$$

When I sent this result to Dr. Sinkov, he replied that he was making a special study of such groups. So we agreed to write consecutive papers, his abstract treatment to follow my geometrical treatment.

Groups of the form $S^l = T^m = (ST)^n = 1$, considered for the sake of analogy

A triangle of angles π/l, π/m, π/n can be drawn on a sphere, or in the euclidean plane, or in the hyperbolic plane, according as the number $1/l + 1/m + 1/n$ is greater than, equal to, or less than unity. By reflecting this triangle in its sides repeatedly, we fill the whole sphere or plane with such triangles, which may be shaded or left white, according to their orientation. Dyck[1] showed that the white (or shaded) triangles correspond to the operators of the abstract group

$$S^l = T^m = (ST)^n = 1.$$

It follows that this group is finite when

$$1/l + 1/m + 1/n > 1,$$

and infinite otherwise. More precisely, its order is

$$\frac{2}{1/l + 1/m + 1/n - 1}$$

whenever this number is positive, and is infinite otherwise. Miller[2] proved that each infinite group has an infinite number of finite factor groups.

Very little is known about the infinite groups, save in the euclidean case

$$1/l + 1/m + 1/n = 1.$$

This case is manageable on account of the presence of self-conjugate subgroups generated by translations, whose quotient groups are obtained by identifying

Received April 11, 1935.

[1] W. Dyck, *Gruppentheoretische Studien*, Math. Ann., vol. 20 (1882), pp. 1–44.

[2] G. A. Miller, *Groups defined by the orders of two generators and the order of their product*, Amer. Jour. of Math., vol. 24 (1902), pp. 96–100.

points of the plane that occupy corresponding positions in a network of period parallelograms. These quotient groups are as follows:[3]

$$S^3 = T^3 = (ST)^3 = (ST^{-1})^b(S^{-1}T)^c = 1, \quad \text{of order } 3(b^2 + bc + c^2);$$

$$S^4 = T^4 = (ST)^2 = (ST^{-1})^b(S^{-1}T)^c = 1, \quad \text{of order } 4(b^2 + c^2) \quad ;$$

$$S^3 = T^6 = (ST)^2 = (TST^{-1}S^{-1})^b(ST^{-1}S^{-1}T)^c = 1, \text{ of order } 6(b^2 + bc + c^2).$$

In particular (putting $b = c$ in the first two cases, and $b = 0$ in the third),

$$S^3 = T^3 = (ST)^3 = (ST^{-1}S^{-1}T)^p = 1 \quad \text{is of order } 9p^2,$$

$$S^4 = T^4 = (ST)^2 = (ST^{-1}S^{-1}T)^p = 1 \quad \text{is of order } 8p^2,$$

$$S^3 = T^6 = (ST)^2 = (ST^{-1}S^{-1}T)^p = 1 \quad \text{is of order } 6p^2.$$

This suggests the following theorem, which is not strictly relevant to our main purpose, but we shall state and prove it, on account of the close analogy with Theorem 4 below (where the geometry is in three dimensions instead of two).

THEOREM 1. *In the group $S^l = T^m = (ST)^n = 1$ with $1/l + 1/m + 1/n \leqq 1$, the commutator of the generators is of infinite period.*[4]

LEMMA 1.1. *On a sphere, or in the euclidean or hyperbolic plane, the continued product of the reflections in the sides of a triangle is an operation that leaves no point invariant.*[5]

Let R_1, R_2, R_3 denote the reflections in the sides of a triangle $A_1A_2A_3$. If possible, let the point P be invariant under the operation $R_1R_2R_3$, so that $P = P \cdot R_1R_2R_3$, i.e., $P \cdot R_3R_2R_1 = P$. If P does not lie in the side A_1A_2, let $P \cdot R_3 = P'$, so that $P' \cdot R_2R_1 = P$. Since R_2R_1 leaves A_3 invariant, $A_3P' = A_3P$. Hence A_3 lies on the perpendicular bisector of PP′, which is A_1A_2 (by definition of P′). On the other hand, if P lies in A_1A_2, we must have $P \cdot R_2R_1 = P$, which makes P coincide with A_3 (or its antipodes). In either case we are led to the absurd conclusion that A_3 lies in A_1A_2. Therefore P cannot exist.

LEMMA 1.2. *For any finite set of (actual) points in the hyperbolic plane, we can define a unique "centroid", which is invariant under all permutations of the points.*

Just as we may represent the points of the elliptic plane by concurrent lines in ordinary space, so also we may represent the points of the hyperbolic plane by time-like lines through a fixed point O of Minkowski three-space. There is, however, one important difference. In the latter case the representative lines are *directed* (in virtue of the "before-after" relation), and so can be replaced by points, equidistant from O, set off along them, either all "before" O or all "after" O. In other words, a "sphere" of time-like radius resembles a hyperboloid of two sheets, and either of the sheets provides a (1, 1) mapping of the

[3] W. Burnside, *Theory of Groups of Finite Order*, Cambridge, 1911, p. 419.

[4] This is a departure from the usual terminology. It seems desirable to speak of the *order* of a group, but the *period* of an operator.

[5] Cf. Theorem 10 of Coxeter, *Discrete groups generated by reflections*, Annals of Math., vol. 35 (1934), p. 602.

hyperbolic plane. Let G denote the centroid of those points of Minkowski space which represent the given set of points of the hyperbolic plane. Then the required centroid of the given points is that point of the hyperbolic plane which is represented by the line OG.

Proof of Theorem 1. Consider the larger group

$$R_1^2 = R_2^2 = R_3^2 = (R_2R_3)^m = (R_3R_1)^n = (R_1R_2)^l = 1, \tag{1.3}$$

in which the given group is a subgroup of index 2, generated by $S = R_1R_2$, $T = R_2R_3$. The generators R_1, R_2, R_3 are reflections in the sides of a triangle of angles π/m, π/n, π/l (in the euclidean or hyperbolic plane, by virtue of the inequality).

By Lemma 1.1, the operation $R_1R_2R_3$ leaves no point invariant. If this operation were of finite period, its powers would transform any given point into a finite set, whose centroid would be invariant. Therefore $R_1R_2R_3$ is of infinite period. But[6] $(R_1R_2R_3)^2 = ST^{-1}S^{-1}T$. Hence[7] $ST^{-1}S^{-1}T$ is of infinite period.

The same result could have been obtained trigonometrically, by showing that the commutator is a rotation through ψ, where[8]

$$\cos^2(\psi/4) = \cos^2\pi/l + \cos^2\pi/m + \cos^2\pi/n + 2\cos\pi/l\cos\pi/m\cos\pi/n.$$

When $1/l + 1/m + 1/n < 1$, ψ is pure imaginary, or rather it is a hyperbolic argument instead of an angle.

By virtue of Theorem 1, we should expect a great variety of factor groups of

$$S^l = T^m = (ST)^n = 1 \qquad (1/l + 1/m + 1/n < 1)$$

to be obtainable by fixing the period of the commutator. Although some progress has been made along these lines,[9] the known results lack generality. There is, however, a general geometrical treatment for the case when we fix the period of the commutator but leave the product (ST) unrestricted. In this respect, the product and commutator exchange rôles in a remarkable manner.

[6] Schwarz used this operation $(R_1R_2R_3)^2$ to determine the triangle of minimum perimeter having one vertex on each side of a given triangle. *Gesammelte math. Abhandlungen*, vol. 2 (1890), p. 344.

[7] We often find it convenient to write the commutator in this form, instead of the orthodox $S^{-1}T^{-1}ST$. In statements of period, this clearly makes no difference, since each of these operators is conjugate to the inverse of the other.

[8] Putting $\theta = 2\pi/l$, $\varphi = 2\pi/m$ and $\cos\lambda = (\cos\pi/l\cos\pi/m + \cos\pi/n)/\sin\pi/l\sin\pi/m$ in formula (1) of G. de B. Robinson, *The real representation of the commutator $S^{-1}T^{-1}ST$ in four dimensions*, Proc. Camb. Phil. Soc., vol. 26 (1930), p. 305.

[9] H. R. Brahana, *Certain perfect groups generated by two operators of orders two and three*, Amer. Jour. of Math., vol. 50 (1928), pp. 345–356.

A. Sinkov, *A set of defining relations for the simple group of order* 1092, Bull. Amer. Math. Soc., vol. 41 (1935), p. 42.

Burnside (op. cit., p. 422) gives the symmetric group of degree 5 in the form $S^2 = T^5 = (ST)^4 = (ST^{-1}ST)^3 = 1$; $S^2 = T^4 = (ST)^5 = (ST^{-1}ST)^2 = 1$ is equally valid.

The case when T is involutory

When $n = 2$, the group (1.3) becomes $[l, m]$, the complete symmetry group of the regular polyhedron[10] $\{l, m\}$. The subgroup generated by R_1R_2 and R_2R_3 is the rotation group $[l, m]'$. When m is even, the operators R_1R_2 and R_3 generate another subgroup (likewise of index 2), which we call $[l', m]$. Writing $S = R_1R_2$, $T = R_3$, we find (since R_1, R_3 are commutative) $S^{-1}TST = (R_2R_3)^2$. Thus the generators of $[l', 2p]$ satisfy

$$S^l = T^2 = (S^{-1}TST)^p = 1. \tag{2.1}$$

They may possibly satisfy other relations, independent of these, but we can assert that $[l', 2p]$ is at least a factor group of (2.1).

To show that $[l', 2p]$ is in fact the whole group (2.1), we observe[11] that the operators $s = S^{-1}$, $s' = TST$ of (2.1) satisfy

$$s^l = s'^l = (ss')^p = 1. \tag{2.2}$$

Since the group (2.2) is invariant under T, it is a subgroup of index 2 in (2.1). Similarly it is of index 2 in

$$s^l = t^2 = (st)^{2p} = 1 \qquad (s' = tst).$$

This last group is $[l, 2p]'$.

Thus (2.2) is of index 4 in $[l, 2p]$, and of index 2 in (2.1). But $[l', 2p]$ is of index 2 in $[l, 2p]$, and is a factor group of (2.1). Therefore $[l', 2p]$ and (2.1) are the same group.

Expressing this result in geometrical terms, we have

THEOREM 2. *The group $[l', 2p]$ has for fundamental region an isosceles triangle of angles $2\pi/l$, $\pi/2p$, $\pi/2p$. It is generated by rotation through $2\pi/l$ about the apex, and reflection in the base.*[12] *Its abstract definition is*

$$S^l = T^2 = (S^{-1}TST)^p = 1.$$

In the figure on page 67, let CAA′ be this isosceles triangle, C′ the image of the apex C in the base AA′, and B the mid-point of AA′ (or of CC′). Then ABC is a fundamental region for $[l, 2p]$, while CAA′ and ACC′ are alternative fundamental regions for $[l, 2p]'$. T is the reflection in AA′, S or s^{-1} is the rotation through $2\pi/l$ about C, s' is the opposite rotation about C′, ss' or $S^{-1}TST$ is the rotation through $2\pi/p$ about A, and t is the rotation through π about B.

By evaluating the side BC of the triangle ABC, we obtain the following

COROLLARY. *The group $[l', 2p]$ is generated by rotation through $2\pi/l$ about a point, and reflection in a line distant λ from this point, where*

$$\sin \pi/l \cos k\lambda = \cos \pi/2p;$$

[10] Bounded by l-gons, m at each vertex.

[11] For this remark I am indebted to Dr. Sinkov.

[12] In other words, this group is generated by rotations about the centers of the faces of the polyhedron $\{l, 2p\}$, and reflections in its edges.

$k = 1, 0$ or i according to the sign of $2/l + 1/p - 1$, the plane being spherical, euclidean, or hyperbolic, in the three cases.

Thus the group is finite only when

$$2/l + 1/p > 1, \tag{2.3}$$

its order then being $4/(2/l + 1/p - 1)$.

$[l', 2]$ is the *equatorial* group C_l^h (occurring in crystallography when $l = 2, 3, 4$ or 6). It is the direct product of the cyclic group of order l with the group of order 2 generated by the equatorial reflection.

$[2', 2p]$ is the *dihedral alternating* group D_p^d. When p is odd, this is the direct product of the dihedral group $[2, p]'$ with the group of order 2 generated by the central inversion. In particular, the *rhombohedral* group $[2', 6]$ can be obtained by adjoining the central inversion to the trigonal dihedral group.

$[3', 4]$ is the *pyritohedral*[13] group T^h. This is the direct product of the tetrahedral group $[3, 3]'$ with the group of order 2 generated by the central inversion.

If $2/l + 1/p = 1$, there are two groups illustrated on pages 66, 67. These are Pólya's[14] D_4^0 and D_3^0, Niggli's[15] $\mathfrak{C}_{4v}^{\text{II}}$ and $\mathfrak{C}_{3v}^{\text{II}}$. In both cases, $(S^{-1}T)^2$ and $(ST)^2$ are translations, and we have the finite factor groups[16]

$$S^4 = T^2 = (S^{-1}TST)^2 = (S^{-1}T)^{2b}(ST)^{2c} = 1, \text{ of order } 8(b^2 + c^2);$$

$$S^3 = T^2 = (S^{-1}TST)^3 = (S^{-1}T)^{2b}(ST)^{2c} = 1, \text{ of order } 6(b^2 + bc + c^2).$$

The subgroups generated by S^{-1} and TST are two of Burnside's groups mentioned above.

The general case

Let $[k_1, k_2, k_3]$ denote the group

$$\begin{aligned} R_1^2 = R_2^2 = R_3^2 = R_4^2 &= (R_1R_2)^{k_1} = (R_2R_3)^{k_2} = (R_3R_4)^{k_3} \\ &= (R_1R_3)^2 = (R_1R_4)^2 = (R_2R_4)^2 = 1. \end{aligned} \tag{3.1}$$

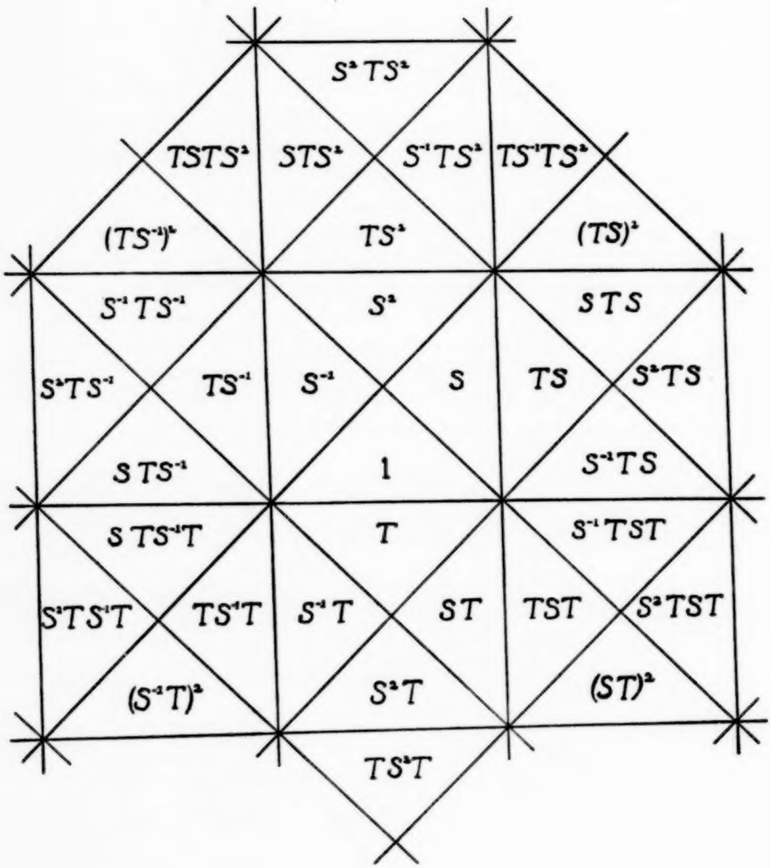
Since every generating relation here involves an even number of generators, there must be a self-conjugate subgroup of index 2, say $[k_1, k_2, k_3]'$, consisting of

[13] A. F. Möbius, *Symmetrische Figuren*, Gesammelte Werke, vol. 2 (1886), p. 672.

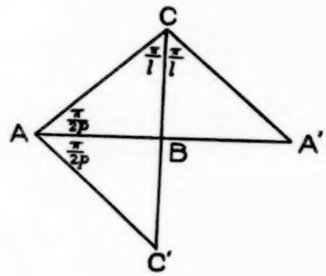
[14] G. Pólya, *Über die Analogie der Kristallsymmetrie in der Ebene*, Zeitschr. für Kristallog., vol. 60 (1924), p. 281.

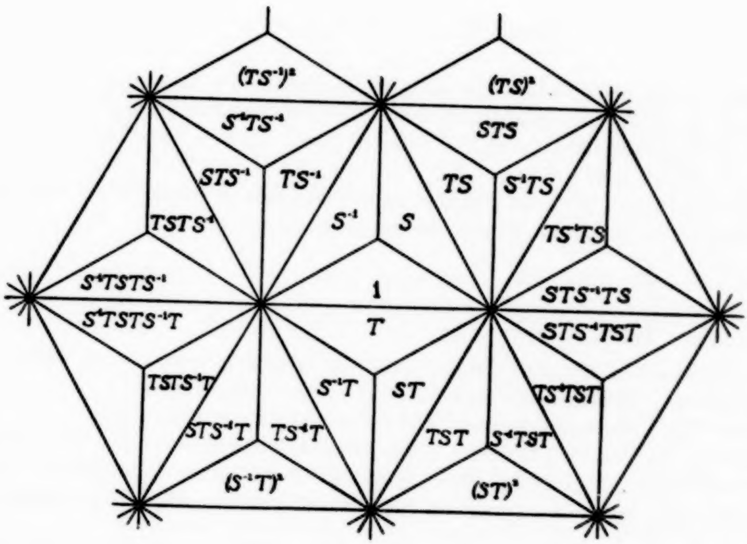
[15] P. Niggli, *Die Flächensymmetrien homogener Diskontinuen*, ibid., p. 291.

[16] These will be studied at greater length in Part II.


S^2TS^2
$TSTS^2$
STS^2
$S^{-1}TS^2$
$TS^{-1}TS^2$
$(TS^{-1})^2$
TS^2
$(TS)^2$
$S^{-1}TS^{-1}$
S^2
STS
S^2TS^{-1}
TS^{-1}
S^{-1}
S
TS
S^2TS
STS^{-1}
1
$S^{-1}TS$
$STS^{-1}T$
T
$S^{-1}TST$
$S^2TS^{-1}T$
$TS^{-1}T$
$S^{-1}T$
ST
TST
S^2TST
$(S^{-1}T)^2$
S^2T
$(ST)^2$
TS^2T

[4',4]


C
π/l
π/l
A
π/2p
π/2p
B
A′
C′


(TS^-1)^2
(TS)^2
S^-1TS^-1
STS
STS^-1
TS^-1
TS
S^-1TS
TSTS^-1
S^-1
S
TS^-1TS
S^-1TSTS^-1
1
STS^-1TS
S^-1TSTS^-1T
T
STS^-1TST
TSTS^-1T
S^-1T
ST
TS^-1TST
STS^-1T
TS^-1T
TST
S^-1TST
(S^-1T)^2
(ST)^2

$[3^+, 6]$

all those operators of $[k_1, k_2, k_3]$ which are products of even numbers of R's. $[k_1, k_2, k_3]'$ is clearly generated by the operators

$$T_1 = R_1R_2, \qquad T_2 = R_2R_3, \qquad T_3 = R_3R_4,$$

which satisfy[17]

$$T_1^{k_1} = T_2^{k_2} = T_3^{k_3} = (T_1T_2)^2 = (T_1T_2T_3)^2 = (T_2T_3)^2 = 1.$$

If k_2 is even, all these relations involve T_2 an even number of times. There must then be a self-conjugate subgroup of index 2, say $[k_1, k_2, k_3]''$, consisting of all those operators of $[k_1, k_2, k_3]'$ which involve T_2 an even number of times. By repeated application of the relations

$$T_2T_1 = T_1^{-1}T_2^{-1}, \qquad T_2T_3 = T_3^{-1}T_2^{-1},$$

$$T_2T_1^{-1} = T_2^2T_1T_2, \qquad T_2T_3^{-1} = T_2^2T_3T_2,$$

all T_2's that occur to an odd power (in the expression for any operator) can be collected in pairs; thus $[k_1, k_2, k_3]''$ is generated by T_1, T_3, and T_2^2.

Since $T_1^{-1}T_3T_1T_3^{-1} = R_2R_1R_3R_4R_1R_2R_4R_3 = R_2R_3R_2R_3 = T_2^2$, the operators T_1 and T_3 generate $[k_1, k_2, k_3]'$ or $[k_1, k_2, k_3]''$ according as k_2 is odd or even.[18] In the latter case they satisfy

$$T_1^{k_1} = T_3^{k_3} = (T_1^{-1}T_3T_1T_3^{-1})^{\frac{1}{2}k_2} = 1.$$

Since we are chiefly concerned with the case when k_2 is even, it is convenient to write $[l, 2p, m]$ instead of $[k_1, k_2, k_3]$. We shall also write S for T_1, and T for T_3^{-1}, so that $S = R_1R_2$, $T = R_4R_3$. Since $[l, 2p, m]''$ and $[m, 2p, l]''$ are identical, we can assume that $l \geqq m$.

Geometrically,[19] $[l, 2p, m]$ is generated by reflections in the faces of a "double-rectangular" tetrahedron, whose dihedral angles are

$$(\mathit{1}\ \ \mathit{2}) = \pi/l, \qquad (\mathit{2}\ \ \mathit{3}) = \pi/2p, \qquad (\mathit{3}\ \ \mathit{4}) = \pi/m,$$

$$(\mathit{1}\ \ \mathit{3}) = \pi/2, \qquad (\mathit{1}\ \ \mathit{4}) = \pi/2, \qquad (\mathit{2}\ \ \mathit{4}) = \pi/2.$$

This tetrahedron will generally have to be in hyperbolic space, but it will be in spherical or euclidean space when l, m, p are sufficiently small.

Since the subgroup $[l, 2p, m]''$ is of index 4, its fundamental region must be made up of four such tetrahedra. From one such tetrahedron, the other three are conveniently derived by reflecting in the faces *1* and *4*. The whole fundamental region is then a tetrahedron in which two opposite edges are perpendicu-

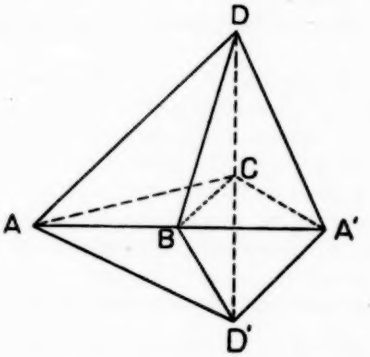
[17] Cf. J. A. Todd, *The groups of symmetries of the regular polytopes*, Proc. Camb. Phil. Soc., vol. 27 (1931), p. 217.

[18] Todd (ibid., p. 229) proved that it is impossible to generate $[3, 4, 3]'$ by two operators.

[19] Todd, ibid., pp. 214, 225.

lar, all the others being equal. Calling the faces *2*, *3*, *2'*, *3'*, the dihedral angles are

$$(2\ \ 2') = 2\pi/l, \qquad (3\ \ 3') = 2\pi/m,$$
$$(2\ \ 3) = (2'\ \ 3) = (2\ \ 3') = (2'\ \ 3') = \pi/2p.$$



In the diagram *1* is BDD′, *2* is ADD′, *2′* is A′DD′, *3* is AA′D, *3′* is AA′D′, *4* is AA′C.

The generators are rotations around the two perpendicular edges DD′, AA′: S carries face *2′* into the position previously occupied by *2*; T carries *3′* into the position previously occupied by *3*. If we associate the original fundamental region with the operator 1, the surrounding regions, beyond faces *2*, *3*, *2′*, *3′*, correspond to the operators S, T, S^{-1}, T^{-1}, respectively. (Face *2* of tetrahedron 1 is face *2′* of tetrahedron S, and so on.)

We have already seen that the rotations S, T satisfy the relations

$$S^l = T^m = (S^{-1}T^{-1}ST)^p = 1, \tag{3.2}$$

and that they suffice to generate the whole group $[l, 2p, m]''$. We shall next prove that *every* relation satisfied by these rotations is a consequence of (3.2).

To find the operator that corresponds to a given region (i.e., that transforms region 1 into the given region), we consider any path from a point within region 1 to a point within the given region. Since we are taking products of operators from left to right,[20] the successive steps along the path have to be written from right to left. (E.g., we pass through face *3* of region 1 into region T, through face *2′* of region T into region $S^{-1}T$, through face *3* of $S^{-1}T$ into $TS^{-1}T$, and through face *3* of $TS^{-1}T$ into $T^2S^{-1}T$.)

Any relation satisfied by the generators provides two different symbols for one region, and so corresponds to a *closed* path. The situation is easily visualized by

[20] Cf. Annals of Math., vol. 35 (1934), p. 599, where, unhappily, I adopted the opposite convention.

thinking of the path as an elastic string, threaded through a network of rigid wires forming the edges of all the tetrahedra. The path can be shrunk to a point by allowing it to slip through these edges, one at a time, the corresponding relation being simplified, at each stage, by means of the generating relation $S^l = 1$ or $T^m = 1$ or $(S^{-1}T^{-1}ST)^p = 1$, according to the type of edge through which the path slips. The sufficiency of these generating relations thus follows from the simple connectivity of the spherical, euclidean, or hyperbolic space, and we have

THEOREM 3. *The group* $[l, 2p, m]''$ *has for fundamental region a tetrahedron with dihedral angles* $2\pi/l$, $2\pi/m$ *at two opposite edges, the four remaining dihedral angles being* $\pi/2p$. *It is generated by rotations about the two special edges.*[21] *Its abstract definition is* $S^l = T^m = (S^{-1}T^{-1}ST)^p = 1$.

The distance between the opposite edges DD′, AA′ of the tetrahedron AA′DD′ is just the length of the edge BC of the double-rectangular tetrahedron ABCD. Evaluating this by spherical trigonometry, we obtain the following

COROLLARY. *The group* $[l, 2p, m]''$ *is generated by rotations through* $2\pi/l$, $2\pi/m$ *about two perpendicular lines, distant* λ *apart, where*

$$\sin \pi/l \sin \pi/m \cos k\lambda = \cos \pi/2p;$$

$k = 1, 0,$ *or* i, *according to the sign of* $\sin \pi/l \sin \pi/m - \cos \pi/2p$, *the space being spherical, euclidean, or hyperbolic, in the three cases.*

Thus the group is finite only when[22]

$$\sin \pi/l \sin \pi/m > \cos \pi/2p. \tag{3.3}$$

When $m = 2$, this condition reduces to (2.3); in fact $[l, 2p, 2]'' \sim [l', 2p]$. When $p = 1$, we have the direct product of cyclic groups of orders l, m: $[l, 2, m]'' \sim [l]' \times [m]'$. The remaining case, when $l = m = 3$ and $p = 2$, appears to be a new discovery: $[3, 4, 3]''$ is of order 288, since $[3, 4, 3]$ is of order 1152.[23] (Actually, it is the commutator subgroup of $[3, 4, 3]$.)

There are no possibilities in the critical case when $\sin \pi/l \sin \pi/m = \cos \pi/2p$, save such as have $m = 2$. When $m = 2$ and $2/l + 1/p = 1$, the fundamental region becomes a baseless prism whose cross-section is the isosceles triangle considered in the two-dimensional representation. In other cases where $m = 2$, the fundamental region has a pair of antipodal vertices (*ideal*, in the hyperbolic case[24]).

The trigonometry involved in proving the above corollary takes no account of

[21] In other words, this group is generated by rotations about the edges of the two reciprocal polytopes $\{m, 2p, l\}$, $\{l, 2p, m\}$. For the theory of infinite regular polytopes, see Coxeter, Proc. Camb. Phil. Soc., vol. 29 (1933), pp. 1–7.

[22] This is, of course, the condition for the polytope $\{l, 2p, m\}$ to be finite.

[23] E. Goursat, *Sur les substitutions orthogonales et les divisions régulières de l'espace*, Ann. Sci. de l'École Norm. Sup., (3), vol. 6 (1889), p. 87.

[24] The fundamental region has two ideal vertices whenever $2/l + 1/p < 1$. All four vertices are ideal if in addition $2/m + 1/p < 1$.

the rationality of l, m, p. We can therefore state that *the commutator of rotations θ and φ about perpendicular lines distant λ apart is a rotation ψ, where*

$$\cos(\psi/4) = \sin\theta/2 \sin\varphi/2 \cos k\lambda.$$

Translating this result (in the spherical case) into terms of euclidean four-space, *the commutator of pure rotations θ and φ about (not absolutely) perpendicular planes inclined at angle λ is a pure rotation ψ, where*[25]

$$\cos(\psi/4) = \sin\theta/2 \sin\varphi/2 \cos\lambda.$$

THEOREM 4.[26] *In the group $S^l = T^m = (S^{-1}T^{-1}ST)^p = 1$ with $\sin\pi/l \sin\pi/m \leqq \cos\pi/2p$, the product of the generators is of infinite period.*

LEMMA 4.1. *In spherical, euclidean, or hyperbolic space, the continued product of the reflections in the faces of a tetrahedron is an operation that leaves no point invariant.*

This is analogous to Lemma 1.1, from which it easily follows.

There is also a precise analogue of Lemma 1.2 in hyperbolic three-space (proved by considering concurrent time-like lines in Minkowski four-space).

Proof of Theorem 4. We have seen that the generators R_1, R_2, R_3, R_4 of $[l, 2p, m]$ are reflections in the faces of a tetrahedron. By Lemma 4.1, the operation $R_1R_2R_4R_3$ $(=ST)$ leaves no point invariant. Hence, the space being euclidean or hyperbolic (in virtue of the inequality), ST is of infinite period.

Clearly this holds also for ST^{-1}.

The same result could have been obtained trigonometrically,[27] by showing that ST (or ST^{-1}) is a double rotation of angles χ, χ', where $\cos^2(\chi/2)$, $\cos^2(\chi'/2)$ are the roots of the equation

$$(x - \cos^2\pi/l)(x - \cos^2\pi/m) = x\cos^2\pi/2p.$$

Elliptic space

THEOREM 5. *Let G denote the dihedral alternating group of any odd degree, or the pyritohedral group, or the new group $[3, 4, 3]''$. Then G has a central of order 2, generated by the central inversion*[28] $(ST)^{\frac{1}{2}h}$, *where h is the period of ST.*

The only case that presents any difficulty is the last. In [3, 4, 3], the central

[25] Cf. G. de B. Robinson, Proc. Camb. Phil. Soc., vol. 26 (1930), p. 309. On replacing ψ by $\psi + 2\pi$, we see that our formula is equivalent to his $\sin^2(\psi/4) = \sin^2(\theta/2)\sin^2(\varphi/2)(1 - P_{12}^2 - P_{34}^2)$ with $P_{12} = 0$ and $P_{34} = \sin\lambda$.

[26] Cf. Theorem 1.

[27] In the notation of F. N. Cole, *On rotations in space of four dimensions*, Amer. Jour. of Math., vol. 12 (1890), pp. 205–208, the product of pure rotations θ and φ about planes $(0, 0, 1, 0, 0, 0)$ and $(P_{23}, P_{31}, 0, P_{14}, P_{24}, P_{34})$ is a double rotation of angles χ, χ', where, since $a = \tan(\theta/2)$, $b = c = f = g = h = \vartheta = \vartheta' = 0$, $D = 1$, $B'' = BB'$, and $\vartheta'' = af'$, $\sec(\chi/2)\sec(\chi'/2) = \sec(\theta/2)\sec(\varphi/2)$, $\tan(\chi/2)\tan(\chi'/2) = P_{34}\tan(\theta/2)\tan(\varphi/2)$. It follows that $\cos(\chi/2)\cos(\chi'/2) = \cos(\theta/2)\cos(\varphi/2)$ and $\cos^2(\chi/2) + \cos^2(\chi'/2) = \cos^2(\theta/2) + \cos^2(\varphi/2) + (1 - P_{34}^2)\sin^2(\theta/2)\sin^2(\varphi/2) = \cos^2(\theta/2) + \cos^2(\varphi/2) + \cos^2(\psi/4)$. For this calculation I am indebted to Dr. Robinson.

[28] Coxeter, Annals of Math., vol. 35 (1934), p. 606.

inversion can be expressed in the form[29] $(R_1R_2R_4R_3)^6$; this is the operator $(ST)^6$ of $[3, 4, 3]''$.

The central quotient groups, $\frac{1}{2}G$, can be regarded as operating in elliptic space. Abstractly, they are given by inserting the extra relation $(ST)^{\frac{1}{2}h} = 1$. Goursat[30] has enumerated all the crystallographic groups in elliptic three-space. Among these we easily pick out XX as the central quotient group of $[3, 4, 3]''$. (It has the right order: 144.)[31]

We thus have the following simple isomorphisms:

$$\tfrac{1}{2}[2, 2p, 2]'' \sim \tfrac{1}{2}[2', 2p] \sim [2, p]' \qquad (p \text{ odd}; h = 2p),$$

$$\tfrac{1}{2}[3, 4, 2]'' \sim \tfrac{1}{2}[3', 4] \sim [3, 3]' \qquad (h = 6),$$

$$\tfrac{1}{2}[3, 4, 3]'' \sim [3, 3]' \times [3, 3]' \qquad (h = 12).$$

An infinite group in which both product and commutator have specified periods

Let $(l, m, n; p)$ denote the group $S^l = T^m = (ST)^n = (ST^{-1}S^{-1}T)^p = 1$. This is not altered by permuting l, m, n, since it can be put into the symmetrical form $S^l = T^m = U^n = STU = (SUT)^p = 1$. After comparing Theorems 1 and 4, it is natural to wonder whether $(l, m, n; p)$ is necessarily finite. We shall show that this is not so, since in fact $(6, 6, 2; 2)$ is infinite.[32]

THEOREM 6. *The group* $S^6 = T^6 = (ST)^2 = (ST^{-1})^{3q} = (ST^{-1}S^{-1}T)^2 = 1$ *is of order* $96\,q^3$.

LEMMA 6.1. *The group* $S^m = T^m = (ST)^2 = (ST^{-1})^n = (ST^{-1}S^{-1}T)^p = 1$ *is a subgroup of index* 2 *in* $s^4 = t^m = (st)^2 = (st^{-1})^{2p} = (st^{-1}s^{-1}t)^n = 1$.

This is easily proved by writing $S = s^2t = st^{-1}s^{-1}$, $T = t^{-1}$, so that $ST = s^2$, $ST^{-1} = st^{-1}s^{-1}t$ and $ST^{-1}S^{-1}T = st^{-1}s^{-1}tt^{-1}s^2t^{-1} = (st^{-1})^2$.

LEMMA 6.2. *The group* $s^4 = t^6 = (st)^2 = (st^{-1})^4 = 1$ *is infinite, and has a representation in euclidean three-space in which the operation* $(st^{-1}s^{-1}t)^3$ *is a translation.*

We know that the group $[4, 3, 4]$ (defined in (3.1)) is the complete symmetry group of the cubic lattice in ordinary space.[33] It has an involutory automorphism R_2', such that

$$R_2 = R_2'R_3R_2', \qquad R_4 = R_2'R_1R_2'.$$

[29] Ibid., p. 608, (vi).

[30] Loc. cit., p. 66.

[31] Dr. Sinkov will clinch the matter by proving abstractly that $\frac{1}{2}[3, 4, 3]''$ is the direct product of two tetrahedral groups. He will also consider other factor groups of $[3, 4, 3]''$, obtained by assigning a smaller period for ST.

[32] This result is of special interest in view of the fact that $(7, 6, 2; 2)$ is finite (of order 2184). Another example is given, in effect, by H. R. Brahana, *On the groups generated by two operators of orders two and three whose product is of order eight*, Amer. Jour. of Math., vol. 53 (1931), p. 901. His results show that $(3, 2, 8; 6)$ is infinite. Elsewhere, we shall prove that there are infinitely many infinite groups $(l, m, n; p)$.

[33] Coxeter, *The densities of the regular polytopes*, Proc. Camb. Phil. Soc., vol. 27 (1931), p. 202 (§3).

(Geometrically, this is the rotation through π about the line joining the points $(\frac{1}{2}, \frac{1}{4}, 0)$ and $(0, \frac{1}{4}, \frac{1}{2})$; it interchanges two reciprocal polytopes $\{4, 3, 4\}$.) By adjoining R_2' to $[4, 3, 4]$ we derive the group

$$R_1^2 = R_2'^2 = R_3^2 = (R_1R_2')^4 = (R_2'R_3)^6 = (R_1R_3)^2 = (R_1R_2'R_3R_2')^4 = 1 .$$

Writing $s = R_1R_2'$, $t = R_2'R_3$, we obtain a subgroup of index 2:

$$s^4 = t^6 = (st)^2 = (st^{-1})^4 = 1 .$$

Clearly $st^{-1}s^{-1}t = (R_1R_2'R_3)^2 = R_1R_2R_4R_3$.

We now make use of Theorem 13 of *Discrete groups generated by reflections*,[34] which tells us that, *of the cycles in which the operation* $R_1R_2R_4R_3$ *of* $[k_1, k_2, k_3]$ *permutes the vertices of the polytope* $\{k_1, k_2, k_3\}$, *one is the cycle of vertices of a Petrie polygon.* Now, the Petrie polygon of the net of cubes $\{4, 3, 4\}$ is a helical polygon, whose sides take the three principal directions in turn, proceeding (say) from the origin to the points (1, 0, 0), (1, 1, 0), (1, 1, 1), (2, 1, 1), (2, 2, 1), $\cdots$. Hence the operation $R_1R_2R_4R_3$ or $st^{-1}s^{-1}t$ is a *trigonal screw*.

Proof of Theorem 6. The translation $(st^{-1}s^{-1}t)^3$ and its conjugates generate a three-dimensional lattice-group. The quotient group $s^4 = t^6 = (st)^2 = (st^{-1})^4 = (st^{-1}s^{-1}t)^3 = 1$ is of order 192 (by direct calculation).[35] By taking a longer translation, we see that the group $s^4 = t^6 = (st)^2 = (st^{-1})^4 = (st^{-1}s^{-1}t)^{3q} = 1$ is of order 192 q^3. The theorem now follows from Lemma 6.1.

In a somewhat similar manner, using the infinite groups

$$\begin{aligned} R_1^2 &= R_2^2 = R_3^2 = R_4^2 = (R_1R_3)^2 = (R_2R_4)^2 \\ &= (R_1R_2)^3 = (R_2R_3)^3 = (R_3R_4)^3 = (R_4R_1)^3 = 1 , \end{aligned}$$

$$s^6 = t^6 = (st)^2 = (st^{-1})^3 = 1 ,$$

we may prove that the group $S^4 = T^6 = (ST)^2 = (ST^{-1})^{4q} = (ST^{-1}S^{-1}T)^3 = 1$ is of order 192 q^3.

Trinity College, Cambridge.

[34] Coxeter, Annals of Math., vol. 35 (1934), p. 605.

[35] By virtue of Lemma 6.1 we need only verify that the subgroup $S^6 = T^6 = (ST)^2 = (ST^{-1})^2 = (ST^{-1}S^{-1}T)^2 = 1$ is of order 96.

THE GROUPS DETERMINED BY THE RELATIONS
$S^l = T^m = (S^{-1}\, T^{-1}\, ST)^p = 1$

Part II

By Abraham Sinkov

1. **Introduction.** The purpose of the second part of this paper is to present a uniform abstract treatment of the spherical and euclidean groups satisfying the conditions given in the title. These have already been considered by Prof. G. A. Miller in four different papers.[1] The methods given here, however, are quite different from those which he used, yielding more general results and a number of properties of the groups not considered by him. In addition, an error is indicated in Miller's results for the case $l, m, p = 3, 3, 2$. He finds the largest group possible under these conditions to be of order 144. It will be shown in what follows (and has already been shown[2] in Part I) that this number should be 288.

2. **Conditions for finiteness.** It can be shown very simply by abstract methods that the only finite groups determined by the relations in question are those given by the solutions of the inequality $\sin \pi/l \sin \pi/m > \cos \pi/2p$. This is accomplished by making use of the known results regarding finiteness in the case of the relations $\sigma^L = \tau^M = (\sigma\tau)^N = 1$. Thus, suppose $p \geqq 3$. Then, since the subgroups $\{S^{-1}, T^{-1}ST\}$ and $\{T^{-1}, S^{-1}TS\}$ correspond to the cases $L, M, N = l, l, p$ and m, m, p respectively, neither l nor m may exceed two. Similarly, if $p = 2$, neither l nor m may exceed three. Hence, except the case $p = 1$, in which the groups are abelian, the only finite groups correspond to the cases $l, m, p = 2, 2, p; 3, 2, 2; 3, 3, 2$. These will be considered first.

3. **The case** $l, m, p = 2, 2, p$. Since the defining relations reduce to $S^2 = T^2 = [(ST)^2]^p = 1$, the period of ST is a divisor of $2p$. It is exactly $2p$ if p is even and either p or $2p$ if p is odd. Hence

Theorem 1. *If p is even, the conditions $S^2 = T^2 = (S^{-1}T^{-1}ST)^p = 1$ determine the dihedral group of order $4p$. If p is odd, these conditions determine either the dihedral group of order $4p$ or that of order $2p$, according as the period of ST is $2p$ or p.*

4. **The case** $l, m, p = 3, 2, 2$. The subgroup $A \equiv \{S^{-1}, T^{-1}ST\}$ is generated by two operators of period three whose product is of period two; it is therefore

Received April 11, 1935.

[1] Proceedings of the National Academy of Sciences, vol. 18 (1932), p. 665; ibid., vol. 19 (1933), p. 199; Tôhoku Mathematical Journal, vol. 38 (1933), p. 1; Journal of the Indian Mathematical Society, vol. 20 (1933), p. 145.

[2] Cf. the discussion of the group [3, 4, 3] following the corollary to Theorem 3.

tetrahedral. Each of its generators is transformed into itself by T. It follows then that G is of order 24 and is the direct product of the tetrahedral group with a group of order 2. This group of order 24 is the most general group satisfying the initial conditions; in it ST is of period 6.

That $(ST)^6 = 1$ is a consequence of $l, m, p = 3, 2, 2$ may be verified rather neatly as follows. The commutator subgroup of G is the non-cyclic group[3] of order 4. Therefore

$$STS^{-1}T\cdot TS^{-1}TS = TS^{-1}TS\cdot STS^{-1}T,\ STSTS = TS^{-1}TS^{-1}TS^{-1}T,\ (ST)^6 = 1\,.$$

As a consequence, in any smaller group which satisfies the conditions $l, m, p = 3, 2, 2$, ST must be of period 3 or 2. If $(ST)^3 = 1$, G is the tetrahedral group. $(ST)^2 = 1$ is impossible, since $S^3 = T^2 = (ST)^2 = 1$ defines the non-cyclic group of order 6, in which the commutators are of period 3.

THEOREM 2. *The conditions* $S^3 = T^2 = (S^{-1}T^{-1}ST)^2 = 1$ *generate either the direct product of the tetrahedral group and the group of order* 2, *or the tetrahedral group, according as* ST *is of period* 6 *or* 3.

5. A general property of different groups corresponding to given l, m, p. The larger group of the two obtained in Theorem 2 is the direct product of the smaller one and a cyclic group. The same statement is true for the two groups obtained in Theorem 1 when p is odd. Such a property occurs in other cases where more than one group results for given l, m, p, and it is of interest to study general conditions under which it occurs. Suppose that, in one of the groups G in question, $S^\alpha T^\beta$ is of period r, and $(S^\alpha T^\beta)^{r/s}$ is invariant in G. The latter operator will generate a cyclic invariant subgroup of order s. If in the quotient group K the operators σ and τ correspond to S and T, then K is defined by the relations $\sigma^l = \tau^m = (\sigma^{-1}\,\tau^{-1}\,\sigma\,\tau)^p = (\sigma^\alpha\,\tau^\beta)^{r/s} = 1$.

We now inquire under what circumstances G will be the direct product of K and a cyclic group of order s. Obviously a necessary and sufficient condition is that it be possible to select one operator from each co-set of G as regards $\{(S^\alpha T^\beta)^{r/s}\}$ in such a way that the totality of operators thus selected will form a group. Now each operator in the co-set containing S is of the form $S(S^\alpha T^\beta)^{ir/s}$; each operator in the co-set containing T is of the form $T(S^\alpha T^\beta)^{jr/s}$. The commutator of two such operators is $S^{-1}T^{-1}ST$. If we let σ_1 and τ_1 represent particular operators in the co-sets containing S and T, they satisfy the relations

$$\sigma_1^l = \tau_1^m = (\sigma_1^{-1}\,\tau_1^{-1}\,\sigma_1\,\tau_1)^p = 1\,,$$

and will generate a group simply isomorphic with K if $(\sigma_1^\alpha\,\tau_1^\beta)^{r/s} = 1$. Now

$$\sigma_1^\alpha\,\tau_1^\beta = (S^\alpha T^\beta)^{1+r(i\alpha+j\beta)/s}\,.$$

[3] A. Sinkov, *A set of defining relations for the simple group of order* 1092, Bull. Amer. Math. Soc., vol. 41 (1935), p. 240.

If it is of period r/s,

$$\frac{r}{s}\left(1 + \frac{r}{s}[i\alpha + j\beta]\right) \equiv 0 \pmod{r},$$

$$1 + \frac{r}{s}(i\alpha + j\beta) \equiv 0 \pmod{s}.$$

This congruence will have a solution if and only if r/s and s are relatively prime. We thus get the following

THEOREM 3. *If in a group G defined by the relations*

$$S^l = T^m = (S^{-1}T^{-1}ST)^p = (S^\alpha T^\beta)^r = 1$$

the operator $(S^\alpha T^\beta)^{r/s}$ is invariant, the quotient group K of G by the cyclic group $\{(S^\alpha T^\beta)^{r/s}\}$ is defined by the relations

$$S^l = T^m = (S^{-1}T^{-1}ST)^p = (S^\alpha T^\beta)^{r/s} = 1.$$

A necessary and sufficient condition that G be the direct product of K and the cyclic group of order s is that r/s and s be relatively prime.

6. **The case** $l, m, p = 3, 3, 2$. The subgroup $A \equiv \{S^{-1}, T^{-1}ST\}$ is generated by two operators of period 3 whose product is of period 2, and is therefore tetrahedral. It is conjugate, under T, to the two subgroups $B \equiv \{S^{-1}, TST^{-1}\}$ and $C \equiv \{T^{-1}ST, TST^{-1}\}$. Hence $D \equiv \{A, B, C\}$ is generated by three operators S^{-1}, $T^{-1}ST$, TST^{-1}, any two of which generate a tetrahedral group. For purposes of convenience, these three generating operators will be designated a, b and c respectively.

It is possible to determine the order of D very readily by enumerating its co-sets with respect to A. In the notation given below, each operator represents a right co-set; for example, c represents all the operators obtained by multiplying each operator of A on the right by c. With this convention, it follows that every operator of D is contained in the eight co-sets $1, c, c^2, cb, c^2a, c^2ab, c^2ab^2, c^2abc$. Numbering these co-sets from 1 to 8, the operators a, b and c are represented as $a = (1)(2, 3, 5)(4, 7, 6)(8)$, $b = (1)(2, 4, 3)(5, 6, 7)(8)$, $c = (1, 2, 3)(4)(5)(6, 8, 7)$. It is thus seen that D is of order 96, if no further restrictions are placed on S and T. Since D is invariant under T, the order of G is at most 288.

This result is at variance with that obtained by Miller,[4] and it is thought advisable to demonstrate it by his method, which helps to establish further properties of the groups.

The powers of the generators S and T give rise to four possible distinct commutators: $\sigma_1 = S^{-1}T^{-1}ST$, $\sigma_2 = TS^{-1}T^{-1}S$, $\sigma_3 = ST^{-1}S^{-1}T$, $\sigma_4 = TST^{-1}S^{-1}$ generating the commutator subgroup[5] H of G. If H contains neither S nor T,

[4] Second reference, footnote 1, p. 200.

[5] G. A. Miller, *On the commutator groups*, Bulletin of the American Mathematical Society, vol. 4 (1897), p. 136.

it yields a quotient group generated by two operators of period 3. This quotient group is abelian, and therefore the index of H under G is at most 9.

Now, as was shown by Miller, σ_1 and σ_4 are each commutative[6] with both σ_2 and σ_3. The subgroup generated by σ_2 and σ_3 is therefore invariant in H, and is of course dihedral. Again $S^{-1}\sigma_2 S = \sigma_2\sigma_4$, $S^{-1}\sigma_3 S = \sigma_1$. Hence $S^{-1}\{\sigma_2, \sigma_3\}S = \{\sigma_2\sigma_4, \sigma_1\}$. This new dihedral group $E \equiv \{\sigma_2\sigma_4, \sigma_1\}$ is also invariant in H. It therefore contains the transform of $\sigma_2\sigma_4$ by σ_3 and the product of this transform by $\sigma_2\sigma_4$, i.e., $\sigma_2\sigma_4 \cdot \sigma_3\sigma_2\sigma_4\sigma_3 = \sigma_2\sigma_3\sigma_2\sigma_3$. If the order of E exceeds 4, the operator $(\sigma_2\sigma_3)^2$ must be contained in its invariant cyclic subgroup. In addition, it is transformed into itself by σ_1. Therefore $\sigma_2\sigma_3$ is at most of period four.[7] Now $\sigma_2\sigma_3 = TS^{-1}T^{-1}S^{-1}T^{-1}S^{-1}T^{-1}T^{-1} = T(S^{-1}T^{-1})^3T^{-1}$, and it follows that the period of ST divides 12.

Suppose $\sigma_2\sigma_3$ is of period 4. Then $\sigma_2\sigma_4\sigma_1$ is also of period 4. But σ_2 and $\sigma_4\sigma_1$ are commutative, and as a result $\sigma_4\sigma_1$ is of period 4. Hence, the two permutable dihedral groups $\{\sigma_1, \sigma_4\}$ and $\{\sigma_2, \sigma_3\}$, which taken together generate H, are each of order 8. Now $\sigma_1\sigma_2\sigma_3\sigma_4 = TS^{-1}TS^{-1}T^{-1}S^{-1}T$ and is of period 2. Therefore $(\sigma_1\sigma_4)^2 = (\sigma_2\sigma_3)^2$. The cross-cut of the two dihedral groups being considered is thus of order 2. Hence H must be of order 32 and G is of order 288.

That this group of order 288 actually exists can be verified by direct calculation.[8] By means of this calculation a representation of G is obtained which is transitive on 32 letters. The subgroup on which this representation is based, i.e., the subgroup which keeps the element 1 unchanged, is $F \equiv \{S, (T^2STST^2)^2\}$; it is a non-cyclic group of order 9.

$$S = (2, 3, 5)\,(4, 7, 11)\,(6, 9, 15)\,(12, 19, 18)\,(8, 13, 21)\,(10, 17, 14)\,(20, 16, 24)\,(27, 32, 30)\,(22, 31, 23)\,(25, 28, 29)\,(1)\,(26)$$

$$T = (1, 2, 4)\,(3, 6, 10)\,(7, 12, 20)\,(5, 8, 14)\,(11, 18, 27)\,(9, 16, 25)\,(19, 17, 26)\,(13, 22, 32)\,(15, 23, 24)\,(21, 30, 29)\,(31)\,(28).$$

This group G of order 288 is the most general group obtainable from the given initial conditions.

We wish next to determine what further groups are possible if additional restrictions are placed on S and T. Obviously these restrictions must be placed on the period of ST and the only possibilities to be considered are the divisors of 12.

Suppose first that $(ST)^6 = 1$. In determining the resulting group, it will first be shown that $(ST)^6$ is invariant in G. For $(\sigma_2\sigma_3)^2 = T(ST)^6T^{-1} = T^{-1}(ST)^6T$ is the only invariant operator in H besides the identity. It is therefore invariant in G. Since it is transformed by T^{-1} into $(ST)^6$, the latter is invariant in G, whose central is consequently of order 2.

[6] Cf. the reference given in note 4, p. 199.

[7] Apparently Miller overlooked this possibility. Cf., also, the first paper mentioned in Note 1, p. 668.

[8] Performed by Dr. Coxeter in connection with his geometric approach.

By Theorem 3, the central quotient group K, of order 144, is defined by the relations $S^3 = T^3 = (S^{-1}T^{-1}ST)^2 = (ST)^6 = 1$. It is not contained in G. In it σ_1, σ_2, σ_3 and σ_4 are all permutable, since $\sigma_1\sigma_4$ and $\sigma_2\sigma_3$ are both of period 2. Its commutator subgroup is therefore an abelian group of order 16 and type (1, 1, 1, 1). The subgroup $U \equiv \{(ST)^2, (TS)^{-2}\}$ is tetrahedral, since $(ST)^2 \cdot (TS)^{-2} = (ST^{-1})^3$, an operator of period 2. The automorphism obtained on replacing S by S^{-1} and leaving T unchanged replaces U by a second tetrahedral group $V \equiv \{(S^{-1}T)^2, (TS^{-1})^2\}$. Each of the generators of V is permutable with both generators of U. Therefore $\{U, V\}$ is the direct product of two tetrahedral groups, and, being of order 144, must coincide with K.

A pair of generating permutations for K is readily obtainable from those for G by adjoining to F the operator $(ST)^6$. The resulting subgroup W which is used as the basis of the new representation is of order 18, and the largest invariant subgroup of G contained in it is of order 2. On obtaining $(ST)^6$ by direct calculation, and equating each pair of elements in each of its cycles, we get

$$S = (1)\ (2, 3, 5)\ (4, 7, 11)\ (6, 9, 15)\ (8, 13, 21)\ (22, 31, 23)$$

$$T = (31)\ (1, 2, 4)\ (3, 6, 11)\ (7, 5, 8)\ (9, 13, 22)\ (15, 23, 21)\,.$$

Suppose $(ST)^4 = 1$. Then, to find the resulting group, it is first necessary to determine the smallest invariant subgroup Q of G which is generated by $(ST)^4$ and its conjugates. Since $(ST)^4$ is transformed by S into $(TS)^4$, Q contains the operators $(ST)^4$ and $(TS)^8 = (S^{-1}T^{-1})^4$. These two operators are of period 3; their product $(ST)^4 \cdot (S^{-1}T^{-1})^4 = (T^{-1}S)^3$ is of period 4. Hence Q is at least[9] of order 24. It follows that the group determined by the relations $S^3 = T^3 = (S^{-1}T^{-1}ST)^2 = (ST)^4 = 1$ is at most of order 12. If we agree that the period of ST must be exactly four, then it is obvious that no group exists for the above set of relations.

In the group of order 144 the operator $(ST)^3$ and its conjugates generate a 4-group. The quotient group, of order 36, is defined by $S^3 = T^3 = (S^{-1}T^{-1}ST)^2 = (ST)^3 = 1$, and is the direct product of the tetrahedral group and a group of order 3.

Finally, if $(ST)^2 = 1$, the resulting group is tetrahedral.

THEOREM 4. *The relations* $S^3 = T^3 = (S^{-1}T^{-1}ST)^2 = 1$ *define only four different groups according as* ST *is of period* 12, 6, 3 *or* 2. *They are of orders* 288, 144, 36, *and* 12, *respectively.*

7. **The case** $l, m, p = 4, 2, 2$. We now wish to consider the euclidean groups, viz., those for which $\sin \pi/l \sin \pi/m = \cos \pi/2p$. The only solutions this equation has are $l, m, p = 4, 2, 2$ and $3, 2, 3$. Let us study the case 4, 2, 2 first. The subgroup $A \equiv \{S^{-1}, T^{-1}ST\}$ is generated by two operators of period 4 whose product is of period 2. These relations $\sigma^4 = \tau^4 = (\sigma\tau)^2 = 1$ were first studied

[9] G. A. Miller, *Groups generated by two operators of order 3 whose product is of order 4*, Bull. Amer. Math. Soc., vol. 26 (1919–20), p. 361–369.

in detail by Burnside,[10] who found that the most general additional restriction that could be imposed on σ and τ is of the form $(\sigma^{-1}\tau)^b\,(\sigma\tau^{-1})^c = 1$, and that the resulting group is of order $4(b^2 + c^2)$. The four relations $\sigma^4 = \tau^4 = (\sigma\tau)^2 = (\sigma^{-1}\tau)^b\,(\sigma\tau^{-1})^c = 1$ imply that the common period of the commutative operators $\sigma\tau^{-1}$ and $\sigma^{-1}\tau$ is $d(b_1^2 + c_1^2)$, where b_1 and c_1 are relatively prime, and $b = db_1$, $c = dc_1$. If we set $\alpha = d(b_1^2 + c_1^2)$, the largest group in which $(\sigma^{-1}\tau)^\alpha = 1$ is of order $4\alpha^2$.

A is invariant under G. For T transforms each of the generators of A into the inverse of the other. Since the adjunction of T to A will generate G, A is of index 2 or 1 under G, according as T is or is not contained in A. Suppose T is in A. Then since the commutator subgroup of A is abelian

$$S^{-1}TST\cdot TSTS^{-1} = TSTS^{-1}\cdot S^{-1}TST\,,$$

$$S^{-1}TS^2TS^2 = TSTS^2\cdot TSTS^{-1} = STS^2T,\quad (TS^2)^4 = 1\,.$$

Now $\sigma^2\tau^2 = TS^2TS^2$. Therefore T is contained in A only when the period of $\sigma^2\tau^2$ divides 2. These exceptional cases will be considered later.

In every other case, A is of index two under G, so that the order of G is $8(b^2 + c^2)$. Since $\sigma^{-1}\tau = (ST)^2$ and $\sigma\tau^{-1} = (S^{-1}T)^2$, the additional relation takes the form $(ST)^{2b}\,(S^{-1}T)^{2c} = 1$. This relation, together with the initial conditions $S^4 = T^2 = (S^{-1}T^{-1}ST)^2 = 1$, implies that the period of $(ST)^2$ is α. If α is even, then ST is of order 2α. If α is odd, there are apparently two possibilities for the period of ST. But one of these leads to a contradiction, for if ST is of odd period, the subgroup A which contains $(ST)^2$ will contain ST, and hence T. This is impossible, since A is of index two under G and it follows that ST is always of even period.

The largest group in which ST is of period 2α is now seen to be of order $8\alpha^2$, and such a group exists for every value of α. A pair of generating permutations for the general group may be obtained as follows. First a pair of permutations[11] are set down which will generate the subgroup A of order $4\alpha^2$:

$$S^{-1} = (1, 2, 3, 4)\,(5, 6, 7, 8)\cdots(4\alpha - 3, 4\alpha - 2, 4\alpha - 1, 4\alpha)$$

$$T^{-1}ST = (3, 4, 5, 6)\,(7, 8, 9, 10)\cdots(4\alpha - 1, 4\alpha, 1, 2)\,.$$

It is now desired to find a permutation T, of period 2, which will transform each of the above substitutions into the inverse of the other. This is relatively simple; if it is supposed only that T replaces 1 by 2, it follows that

$$T = (1, 2)\,(3, 4\alpha)\,(4, 4\alpha - 1)\cdots(2\alpha + 1, 2\alpha + 2)\,.$$

This permutation, together with

$$S = (1, 4, 3, 2)\,(5, 8, 7, 6)\cdots(4\alpha - 3, 4\alpha, 4\alpha - 1, 4\alpha - 2)\,,$$

[10] W. Burnside, *Theory of Groups of Finite Order*, Cambridge, 1911, p. 416.

[11] W. E. Edington, *Abstract group definitions and applications*, Transactions Amer. Math. Soc., vol. 25 (1923), p. 198.

generates the required group of order $8\alpha^2$. It is easy to verify that $S^{-1}T^{-1}ST$ is of period 2 and that ST is of period 2α.

Consider now this group G of order $8\alpha^2$. Let P be an operator in the group of the form $(ST)^{2b}(S^{-1}T)^{2c}$, where $b_1^2 + c_1^2$ divides α. Then the period of P is $b_1^2 + c_1^2$. If we form the invariant subgroup U generated by the complete set of conjugates containing P, and then determine the quotient group V of G with respect to U, then V is defined by the relations $S^4 = T^2 = (S^{-1}T^{-1}ST)^2 = (ST)^{2b}(S^{-1}T)^{2c} = 1$, which imply $(ST)^{2\alpha} = 1$. These relations are known to yield a group of order $8(b^2 + c^2)$, which is the quotient of $8\alpha^2$ by $b_1^2 + c_1^2$. Therefore U is the cyclic group generated by P and we have the following

THEOREM 5. *In the group of order $8\alpha^2$ defined by the relations $l, m, p = 4, 2, 2$ and the additional condition $(ST)^{2\alpha} = 1$, any operator of the form $(ST)^{2b}(S^{-1}T)^{2c}$ for which $b_1^2 + c_1^2$ divides α is conjugate to powers of itself only.*

Since the quotient group of G by $\{P\}$ is of order $8(b^2 + c^2)$, it follows from the existence of a group G for every α that a group K of order $8(b^2 + c^2)$ exists for every pair of numbers b and c.

To complete the investigation, it is now necessary to consider the exceptional cases $(\sigma^2\tau^2)^2 = 1$. In each of these cases, it is possible to select two operators, viz. σ^{-1} and $\sigma\tau$, which are of period 4 and 2 respectively, have a commutator of period 2 and generate the entire group. As a result, the exceptional cases are found to coincide with the groups of order $8\alpha^2$ and $8(h^2 + h^2)$ for $\alpha = 1$ and $h = 1$.

In any group K the operator $(ST)^{2d}$ can be shown to be conjugate to its own powers only. Hence, the quotient group of K by $\{(ST)^{2d}\}$ is of order $8d^2$ and is defined by the relations $S^4 = T^2 = (S^{-1}T^{-1}ST)^2 = (ST)^{2d} = 1$. Since it is one of the groups G, there exists a reciprocal relation between the groups G and the groups K, which is best expressed as follows. For every number w of the form $b_1^2 + c_1^2$ there exists an infinite family of groups K of order $8d^2(b_1^2 + c_1^2)$ which has with the groups G the reciprocal relation that a group in either family is obtainable as a quotient group in the other.

Let us consider the special case when K is the central quotient group of G. In order to do this we first determine the invariant operators of G. Since

$$T^{-1}(ST)^{2b}(S^{-1}T)^{2c}T = (TS)^{2b}(TS^{-1})^{2c} = (ST)^{-2c}(S^{-1}T)^{-2b},$$

a necessary condition is given by $b = c = \alpha/2$. This implies, of course, that α is even. Now

$$[(ST)^2]^{\alpha/2}[(S^{-1}T)^2]^{\alpha/2} = (S^2T)^{\alpha} = (\sigma^2\tau^2)^{\alpha/2}$$

is known to be the only invariant operator in A. Since G may be generated by adjoining S^2T to A, it follows that the condition is also sufficient. The central of G is therefore of order 2 provided α is even, and in that case the central quotient group K is of order $16(\alpha/2)^2 = 4\alpha^2$. If we set $\alpha = 2h$, there is a group of order $16h^2$ for every h.

Now the group G of order $8h^2$ is obtained from the above group by taking the quotient group with respect to $\{(ST)^{2h}\}$. Obviously, then, $(ST)^{2h}$ is invariant in K. Furthermore, the central of any group G is at most of order 2. Therefore the central of K is at most of order 4. It will be of order 4 only if the two operators $(S^2T)^h$, $(ST)^{2h}\,(S^2T)^h$ are separately invariant. But $S^{-1}(S^2T)^2S = (ST)^4(TS^2)^2$, so that $S^{-1}(S^2T)^hS = (ST)^{2h}\,(S^2T)^h$. The central of K is therefore of order two, and the groups G are also central quotient groups. It is thus seen that the reciprocal property previously mentioned becomes in this special case a reciprocal relationship between central quotient groups.[12]

It is a consequence of what has been demonstrated up to this point that the period of ST in any group defined by the relations $S^4 = T^2 = (S^{-1}T^{-1}ST)^2 = 1$ is an even number. The assumption of an odd number for the period of ST requires A to coincide with G, and hence leads to a contradiction, so that there is no group to correspond to the relations $S^4 = T^2 = (ST)^{2r+1} = (S^{-1}T^{-1}ST)^2 = 1$.

THEOREM 6. *The most general relation that may be adjoined to conditions* $S^4 = T^2 = (S^{-1}T^{-1}ST)^2 = 1$ *is* $(ST)^{2b}\,(S^{-1}T)^{2c} = 1$. *The four relations define a group of order* $8(b^2 + c^2)$, *and such a group exists for every pair of numbers* b *and* c. *In any one of these groups the period of* ST *is* $2d(b_1^2 + c_1^2)$, *where* b_1 *and* c_1 *are relatively prime and* $b = db_1$; $c = dc_1$. *No group is possible in which the period of* ST *is an odd number.*

A procedure for obtaining a pair of generating permutations for any one of the groups of order $8(b^2 + c^2)$ from those already given for the groups G of order $8\alpha^2$ will now be outlined, the case $b = c$ being selected for purposes of illustration. The representation obtained for G is transitive and of degree 4α. Therefore, the subgroup B which keeps the element 1 unchanged is of order 2α. It is non-invariant in G and involves no invariant subgroup of G. Since $S^{-1}T$ is of period 2α and keeps the element 1 unchanged, $B \equiv \{S^{-1}T\}$. Suppose now that the operator $(ST)^\alpha(S^{-1}T)^\alpha$ is adjoined to B, yielding a subgroup C of order 4α. If the largest invariant subgroup of G contained in C is of order two, the central quotient group K of G will be obtained by using C as the basis of a new transitive representation. The actual mechanics would involve the calculation of $(ST)^\alpha(S^{-1}T)^\alpha$ and the subsequent equating in both S and T of each pair of elements which occur in the individual cycles of $(ST)^\alpha(S^{-1}T)^\alpha$.

Now it happens that C involves an invariant subgroup of order four, viz., $\{(ST)^\alpha, (S^{-1}T)^\alpha\}$. Moreover, this is the largest invariant subgroup of G contained in C. Hence, in order to get a representation for K, it is necessary to obtain first a representation of G, in which the basic group B is replaced by a subgroup B_1, which does not contain $(S^{-1}T)^\alpha$. If α is divisible by 2^p but not by 2^{p+1}, this may be accomplished by setting $B_1 \equiv \{(S^{-1}T)^\beta\}$, where $\beta = 2^{p+1}$. When this has been done, the new representation for G will involve $2^{p+2}\alpha$ let-

[12] These two infinite families of order $8\alpha^2$ and $16h^2$, corresponding to the cases $c = 0$ and $b = c$, are the two families obtained by Miller in his study of the same case.

ters. To get the corresponding generators, we replace each cycle (a, b) of T by the 2^{p+1} cycles

$$(a, b)\ (4\alpha + a, 4\alpha + b) \cdots ([2^{p+3} - 4]\ \alpha + a, [2^{p+3} - 4]\ \alpha + b)\,.$$

Each cycle $(4k - 3, 4k, 4k - 1, 4k - 2)$ of S is replaced by the 2^{p+1} cycles

$$(4k - 3, 4\alpha + 4k, 4k - 1, 4\alpha + 4k - 2)\ (4\alpha + 4k - 3, 8\alpha + 4k\,,$$
$$4\alpha + 4k - 1, 8\alpha + 4k - 2) \cdots ([2^{p+3} - 4]\alpha + 4k - 3, 4k\,,$$
$$[2^{p+3} - 4]\ \alpha + 4k - 1, 4k - 2)\,.$$

The operator $(ST)^{\alpha}(S^{-1}T)^{\alpha}$ is now found to be

$$(1, 2\alpha + 1)\ (2, 2\alpha + 2) \cdots (2\alpha, 4\alpha)\ (4\alpha + 1, 6\alpha + 1) \cdots (6\alpha, 8\alpha)$$
$$\cdots ([2^{p+3} - 4]\ \alpha + 1, [2^{p+3} - 2]\ \alpha + 1) \cdots ([2^{p+3} - 2]\ \alpha, 2^{p+3}\alpha)\,.$$

Equating each of these pairs of numbers in the new forms just obtained for S and T, we get a representation for K which is transitive on $2^{p+2}\alpha$ letters.

In general, if α is divisible by $(b_1^2 + c_1^2)^p$ but not by $(b_1^2 + c_1^2)^{p+1}$, the above procedure will yield a transitive representation on $4(b_1^2 + c_1^2)^p\alpha$ letters for the group of order $8(b^2 + c^2)$.

8. **The case** $l, m, p = 3, 2, 3$. The subgroup $A \equiv \{S^{-1}, T^{-1}ST\}$ is generated by two operators of period 3 whose product is of period 3. These relations $\sigma^3 = \tau^3 = (\sigma\tau)^3 = 1$ were also studied by Burnside.[13] The most general additional relation that can be imposed on σ and τ is of the form $(\sigma^{-1}\tau)^b\,(\sigma\tau^{-1})^c = 1$, and the resulting group is of order $3(b^2 + bc + c^2)$. The four relations imply that the common period of the commutative operators $\sigma^{-1}\tau$ and $\sigma\tau^{-1}$ is $d(b_1^2 + b_1c_1 + c_1^2)$, where d, b_1 and c_1 have the same meaning as before.

A is invariant in G and of index 2 or 1 according as T is or is not contained in A. Suppose T is in A. Then, since the commutator subgroup of A is abelian,

$$S^{-1}TST\cdot TSTS^{-1} = TSTS^{-1}\cdot S^{-1}TST,\ S^{-1}TS^{-1}TS^{-1} = TSTSTST,\ (ST)^6 = 1\,.$$

But $\sigma^{-1}\tau = (ST)^2$; therefore $(\sigma^{-1}\tau)^3 = 1$. Setting these cases aside for the moment, we see that G is of order $6(b^2 + bc + c^2)$. In it the additional relation is of the form $(ST)^{2b}\,(S^{-1}T)^{2c} = 1$.

The treatment from this point follows the same lines as the case $l, m, p = 4, 2, 2$. The additional relation implies that ST is of period 2α; the largest group G in which $(ST)^{2\alpha} = 1$ is of order $6\alpha^2$. Such a group exists for every value of α and a pair of generating permutations can be set up for the general case.

$$S = (1, 3, 2)\ (4, 6, 5)\ (7, 9, 8) \cdots (3\alpha - 2, 3\alpha, 3\alpha - 1)$$

$$T = (2, 3\alpha)\ (3, 3\alpha - 1)\ (4, 3\alpha - 2) \cdots \left(\left[\frac{3\alpha}{2}\right] + 1, 3\alpha + 2 - \left[\frac{3\alpha}{2}\right]\right).$$

[13] Burnside, loc. cit., p. 414.

In any such group G, an operator $P \equiv (ST)^{2b} (S^{-1}T)^{2c}$, for which $b_1^2 + b_1c_1 + c_1^2$ divides α, is conjugate to its own powers only, and therefore a group K of order $6(b^2 + bc + c^2)$ exists for every pair of numbers b and c. A pair of generating permutations for K can be obtained from those for G. The representation is transitive and involves $3(b_1^2 + b_1c_1 + c_1^2)^{p-1}\alpha$ letters whenever α is divisible by $(b_1^2 + b_1c_1 + c_1^2)^p$ but not by $(b_1^2 + b_1c_1 + c_1^2)^{p+1}$. The exceptional cases when A and G coincide are found to be contained in the groups of order $6(b^2 + bc + c^2)$.

The groups G are obtainable from the groups K as quotient groups with respect to the cyclic group $\{(ST)^{2d}\}$, so that the same kind of reciprocal relationship is obtained. However, the property of reciprocal central quotient groups does not enter. For $c = 0$ and $b = c$ we have two infinite families of groups,[14] G of order $6\alpha^2$ and K of order $18h^2$. The former are central quotient groups of the latter, but the converse is not true. The operator $(ST)^{2\alpha/3} (S^{-1}T)^{2\alpha/3}$ is conjugate to its inverse in G.

Theorem 7. *The most general relation that may be adjoined to the conditions* $S^3 = T^2 = (S^{-1}T^{-1}ST)^3 = 1$ *is* $(ST)^{2b} (S^{-1}T)^{2c} = 1$. *The four relations define a group of order* $6(b^2 + bc + c^2)$, *and such a group exists for every pair of numbers* b *and* c. *In any one of these groups the period of* ST *is* $2d(b_1^2 + b_1c_1 + c_1^2)$; *no group is possible in which the period of* ST *is an odd number.*

Washington, D. C.

[14] Here again, these two infinite families, corresponding to the special cases $c = 0$ and $b = c$, are the only groups obtained by Miller in his study.

EXTENSIONS OF THEOREMS OF DESCARTES AND LAGUERRE TO THE COMPLEX DOMAIN

By I. J. Schoenberg

§1. Introduction and statement of results

1. Much attention has been devoted to the important problem of finding limitations for the absolute values of the zeros of a polynomial in terms of the absolute values of the coefficients of the polynomial. Much less is known about the arguments of the zeros when only the arguments of the coefficients are taken into account. Regarding this latter problem in its full generality, I can mention only an interesting article by A. J. Kempner.[1] The classical rule of Descartes is a contribution to this problem for real equations as far as real roots are concerned. Obreschkoff's extension of this rule (Theorem I below)[2] to those roots of real equations which lie in a certain angular neighborhood of the real axis points the way to a new extension of Descartes' rule which will take care of all the real or complex roots of real or complex equations (Theorem II and subsequent remarks in section 4). Furthermore, it will be shown that a theorem of Laguerre (Theorem V) actually applies to the roots in rather extended domains of the complex plane (Theorem VI). All these results are derived by means of the fruitful idea used by Obreschkoff in proving his Theorem I. It consists in letting the original theorems (of Descartes and Laguerre respectively) extend themselves, so to speak, to complex roots in certain domains by means of a classical theorem of Cauchy applied along the boundary of the corresponding domain.

2. The following extension of the rule of Descartes is due to N. Obreschkoff (loc. cit.).

Theorem I. *Let*

$$f(x) \equiv a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0 \tag{1}$$

Received August 24, 1934, by the Editors of the Annals of Mathematics, accepted by them, and later transferred to this journal.

[1] A. J. Kempner, *Über die Separation complexer Wurzeln algebraischer Gleichungen*, Math. Annalen, vol. 85 (1922), pp. 49–59. Using systematically the fact that a sum of complex numbers with positive real parts can not vanish, Kempner shows how to divide the entire plane into consecutive infinite sectors S_1, S_1', S_2, S_2', S_3, S_3', $\cdots$ with a common vertex at the origin such that there are no zeros within the sectors S_1, S_2, S_3, $\cdots$, while each of the sectors S_1', S_2', S_3', $\cdots$ contains at least one zero of the given polynomial. See footnote 3 below. See also his recent comprehensive article *On the complex roots of algebraic equations*, Bull. Amer. Math. Soc., vol. 41 (1935), pp. 809–843.

[2] N. Obreschkoff, *Sur un problème de Laguerre*, Comptes rendus de l'Acad. des Sciences, vol. 177 (1923), pp. 102–104.

be an algebraic equation with real coefficients. The number v_a *of variations in the sequence of its coefficients is not merely an upper bound for the number* $Z(x > 0)$ *of its positive roots (rule of Descartes), but also an upper bound for the number of all of its roots which lie inside the sector* $|\arg x| < \pi/n$, *i.e.,*

$$Z\{|\arg x| < \pi/n\} \leqq v_a. \tag{2}$$

Substituting $x = -z$ into (1) we also find an estimate for the number of roots within the opposite sector $\pi - (\pi/n) < \arg x < \pi + (\pi/n)$. What about Z for some sector with the vertex 0 and containing, say, the upper half of the imaginary axis in its interior? If we try to use Theorem I after performing the rotation $x = iz$, we shall find in general a new equation in z with *complex* coefficients to which Theorem I does not apply.

3. This remark suggests the desirability of extending Obreschkoff's theorem to equations (1) with complex coefficients. This obviously necessitates an extension of v_a to complex numbers $a_0, a_1, \cdots, a_n$. A natural extension of this notion is as follows. Let us mark in the complex a-plane the points $a_0, a_1, \cdots, a_n$. Through the origin O of the a-plane we draw any two straight lines Δ and Δ' dividing the plane into four consecutive sectors A, B, C, D with the following property: *one* of the two pairs of opposite sectors, A and C, say, shall not contain any of the points a_ν in its interior. Call ψ (we assume $\psi \neq 0$, hence $0 < \psi \leqq \pi$) the common aperture of A and C. The points a_ν are thus assumed to lie inside or on the boundary of the remaining opposite sectors B and D. For this reason we say that the angles A and C form a *separating double sector* (of aperture ψ) for the coefficients a_ν. In the process of going through the sequence of points $a_0, a_1, \cdots, a_n$ we count the number of times we have to pass from the sector B to D or vice-versa, and call this number $v_a(S)$ [= *number of variations of our sequence with respect to the separating double sector* $S = (A, C)$]. In counting these variations, points a_ν which are at the origin O are to be disregarded. If all the a_ν are real, we can take $\psi = \pi$ and thus get the usual number v_a.

With this definition we can now state

THEOREM II. *Let* (1) *be an algebraic equation with real or complex coefficients. In the complex a-plane mark the points* a_ν *and draw a separating double sector* S *of aperture* ψ $(0 < \psi \leqq \pi)$. *Then*

$$Z\{|\arg x| < \psi/n\} \leqq v_a(S). \tag{3}$$

For ψ *and* n *fixed the sector* $|\arg x| < \psi/n$ *is the largest domain for which* (3) *always holds, for this inequality may fail to hold if we add to the sector even a single point on its boundary.*

If the equation (1) is real, we can take $\psi = \pi$, and (3) reduces to Obreschkoff's inequality (2).

4. Now if we wish to find an estimate for the number of roots within some sector whose bisector makes with the positive axis the angle θ, we substitute

$x = e^{i\theta}z$ into (1) and denote by S' a separating double sector of aperture ψ' for the coefficients a'_ν of the transformed equation

$$(4)\quad f(e^{i\theta}z) \equiv a_0 + a_1 e^{i\theta}z + \cdots + a_n e^{in\theta}z^n \equiv a'_0 + a'_1 z + \cdots + a'_n z^n = 0 .$$

We obtain by Theorem II[3]

$$Z\left\{\theta - \frac{\psi'}{n} < \arg x < \theta + \frac{\psi'}{n}\right\} \leqq v_{a'}(S').$$

Thus Theorem II gives information about the distribution of the arguments of the roots of (1) by means of the arguments of the coefficients of this equation. Additional information may be obtained by changing the origin of the x-plane.

In Theorem II it will in general be possible to find for the same equation (1) a considerable number of separating double sectors S with different $v_a(S)$ corresponding to them. It is desirable, in order to make the inequality (3) more effective, to have ψ as large and $v_a(S)$ as small as possible. Disregarding $v_a(S)$ for the moment, we can always choose ψ so as to exceed a certain constant due to the following elementary theorem.

THEOREM III. *It is always possible to find for the coefficients of an equation* (1) *of degree n a separating double sector of aperture $\psi \geqq \pi/(n+1)$. The constant $\pi/(n+1)$ is here the largest possible for a given degree n.*

Hence Theorem II will give an upper bound for the number of the roots of (1) within any sector with vertex at the origin and of aperture $2\pi/[n(n+1)]$, for any rotation $x = e^{i\theta}z$ will give a new equation (4) to which a separating sector of aperture $\geqq \pi/(n+1)$ can be found according to Theorem III.

The difference of the two sides of the inequality (2) is in any case an even number. Easy examples will show that this is not always true for the general inequality (3).

5. Laguerre has extended Descartes' theorem to exponential polynomials

$$(5)\quad F(x) \equiv a_0 e^{\lambda_0 x} + a_1 e^{\lambda_1 x} + \cdots + a_n e^{\lambda_n x} = 0 \qquad (\lambda_0 < \lambda_1 < \cdots < \lambda_n),$$

with $a_\nu \gtrless 0$, and found that $Z\{-\infty < x < \infty\} \leqq v_a$. An extension of this result similar to Theorem II follows.

THEOREM IV. *Let* (5) *be an exponential equation with real or complex coefficients and real monotonically increasing exponents. In the complex a-plane mark the points a_ν and draw a separating double sector S of aperture ψ $(0 < \psi \leqq \pi)$. Then*

$$(6)\qquad Z\{\,|\mathfrak{J}x| < \psi/(\lambda_n - \lambda_0)\} \leqq v_a(S).$$

[3] A "planetarium" as suggested by Kempner (loc. cit., p. 54) would greatly facilitate the determination of S', ψ' and $V_{a'}(S')$ for arbitrary θ in any numerical case. It is a watch-like instrument with n hands (vectors) capable of rotating with angular velocities of ratios $1:2:3:\cdots:n$ and capable of starting from arbitrarily assigned positions (corresponding to the arguments of $a_1, a_2, \cdots, a_n$ if $a_0 = 1$). Thus an ordinary watch whose hands can be moved into any initial positions will take care of any equations of the type $1 + a_1 x + a_{12}x^{12} = 0$.

Note that Theorem II is a special case of Theorem IV for $\lambda_\nu = \nu$ and the new variable $z = e^x$. M. Marden[4] has previously shown that $|\Im x| < \psi/(\lambda_n - \lambda_0)$ is a zero-free region of $F(x)$ provided $v_a(S) = 0$. Just as Theorem II could be applied to various angles by rotation of the x-plane, so can Theorem IV be applied to various horizontal strips by vertical translation of the x-plane. Theorem III insures the applicability of Theorem IV to any horizontal strip[5] of at least the width $2\pi/[(n+1)(\lambda_n - \lambda_0)]$.

It is of interest to note that the constant $\psi/(\lambda_n - \lambda_0)$ of Theorem IV is a function of the difference $\lambda_n - \lambda_0$, and hence independent of the number of terms of the exponential sum (6). For this reason it can readily be extended to integral functions of the type

$$F(x) = \int_{\lambda_0}^{\lambda_1} e^{\lambda x}\varphi(\lambda)\, d\lambda .$$

6. Let us consider now a rational function of the form

$$(7)\quad F(x) \equiv \frac{a_0}{x-\alpha_0} + \frac{a_1}{x-\alpha_1} + \cdots + \frac{a_n}{x-\alpha_n} \quad (\alpha_0 > \alpha_1 > \cdots > \alpha_n;\, a_\nu \text{ real} \neq 0).$$

It is well known that if a_ν and $a_{\nu+1}$ have the same sign, then $F(x)$ has an odd number of zeros inside the interval $(\alpha_{\nu+1}, \alpha_\nu)$. Denoting by v_a the number of variations in the sequence $a_0, a_1, \cdots, a_n$, we conclude that $F(x)$ has at least $n - v_a$ zeros in $\alpha_n < x < \alpha_0$ and therefore at most v_a zeros in the complex plane outside the interval $\alpha_n < x < \alpha_0$; in particular, we thus get $Z\{\alpha_0 < x < \infty\} \leqq v_a$. Let us denote by $v(a_0 + a_1 + \cdots + a_n)$ (= *the number of variations in the sum* $a_0 + a_1 + \cdots + a_n$) the number of ordinary variations in the sequence of partial sums $a_0, a_0 + a_1, a_0 + a_1 + a_2, \cdots, a_0 + \cdots + a_n$. A better estimate of $Z\{\alpha_0 < x < \infty\}$ is furnished by the following theorem of Laguerre.[6]

[4] Morris Marden, *On the zeros of certain rational functions*, Trans. Amer. Math. Soc., vol. 32 (1930), p. 662.

[5] A theorem on more general exponential polynomials (a_ν = a polynomial in x) proposed by G. Pólya as a problem and proved by Obreschkoff, Jahresbericht der Deutschen Math.-Vereinigung, vol. 37 (1928), pp. 82–83, Lösung der Aufgabe 24, when specialized to polynomials of type (5) with $a_0 a_n \neq 0$ gives for any finite interval (α, β) the interesting inequalities

$$(6')\qquad \frac{(\lambda_n - \lambda_0)(\beta - \alpha)}{2\pi} - n \leqq Z\{\alpha \leqq \Im x \leqq \beta\} \leqq \frac{(\lambda_n - \lambda_0)(\beta - \alpha)}{2\pi} + n.$$

Results (6) and (6′) do not imply one another. If we apply (6′) to the case of ordinary polynomials $f(z) = a_0 + a_1 z + \cdots + a_n z^n$ by writing $\lambda_\nu = \nu$ $(\nu = 0, 1, \cdots, n)$, $z = e^x$, $0 < \beta - \alpha < 2\pi$, we get the result

$$n\left(\frac{\beta-\alpha}{2\pi} - 1\right) \leqq Z\{\alpha \leqq \arg z \leqq \beta\} \leqq n\left(\frac{\beta-\alpha}{2\pi} + 1\right),$$

which is trivial, for the lower bound is < 0 while the upper bound exceeds n. This was to be expected, for (6′) does not utilize the arguments of the coefficients of (5). I am indebted to Harry Matison for calling my attention to Pólya's theorem.

[6] E. Laguerre, *Oeuvres*, vol. 1, p. 41. See footnote 7 below.

THEOREM V. *The number of real zeros of the function $F(x)$ defined by* (7) *which are greater than α_0 does not exceed the number of variations in the sum $a_0 + a_1 + \cdots + a_n$.*

Moreover, if $\sum_0^n a_\nu \neq 0$, the difference $v(a_0 + \cdots + a_n) - Z\{\alpha_0 < x < \infty\}$ is an even number. Our extension of this theorem to the complex roots of (7) is as follows.

THEOREM VI. *If $\sum_0^n a_\nu \neq 0$, in Laguerre's Theorem V $v(a_0 + \cdots + a_n)$ is not merely an upper bound*[7] *for the number $Z\{\alpha_0 < x < \infty\}$ of real zeros of* (7) *which are greater than α_0, but also an upper bound for the number of its zeros of real part $\geqq \alpha_0$, i.e.,*

$$Z\{\Re x \geqq \alpha_0\} < v(a_0 + a_1 + \cdots + a_n). \tag{8}$$

If $\sum_0^n a_\nu = 0$, the same inequality holds for $Z\{\Re x > \alpha_0\}$ instead of $Z\{\Re x \geqq \alpha_0\}$.

The property concerning the evenness of the difference remains valid also in the half-plane (provided $\sum_0^n a_\nu \neq 0$), for possible complex zeros appear in pairs.

By means of a simple homographic transformation, Laguerre (loc. cit., pp. 42–47) derived from Theorem V the following elegant estimate for the number of zeros of (7) within the interval $\xi < x < \alpha_\nu$ $(\alpha_{\nu+1} < \xi < \alpha_\nu)$:

$$\begin{aligned} Z\{\xi < x < \alpha_\nu\} \leqq v\biggl(&\frac{a_\nu}{\xi - \alpha_\nu} + \frac{a_{\nu-1}}{\xi - \alpha_{\nu-1}} + \cdots + \frac{a_0}{\xi - \alpha_0} + \frac{a_n}{\xi - \alpha_n} \\ &+ \frac{a_{n-1}}{\xi - \alpha_{n-1}} + \cdots + \frac{a_{\nu+1}}{\xi - \alpha_{\nu+1}}\biggr). \end{aligned} \tag{9}$$

Similarly Theorem VI shows that the right side of (9) is an upper bound for Z within the circle with (ξ, α_ν) as diameter and, if $F(\xi) \neq 0$, we may even count the zeros on the boundary of this circle.

7. Laguerre's result in its extended form is of importance, because of the ease with which the quotient of a polynomial $f(x)$ of degree n by $(x - \alpha_0)(x - \alpha_1) \cdots (x - \alpha_n)$ can be decomposed into partial fractions and thus put in the form (7). Thus from the formula

$$\frac{f(x)}{(x-a)(x-a-h)\cdots(x-a-nh)} = \frac{(-1)^n}{n!\,h^n} \sum_{\nu=0}^{n} (-1)^\nu \binom{n}{\nu} \frac{f(a+\nu h)}{x - a - \nu h}$$

[7] The inequality $v_a \leqq v(a_0 + \cdots + a_n)$ shows that Laguerre's upper bound for $Z\{\alpha_0 < x < \infty\}$ is more accurate than the upper bound v_a derived at the beginning of section 6. However, the weaker (i.e., larger) bound v_a holds for all zeros of (7) outside the real interval (α_n, α_0), while Laguerre's bound holds for the zeros in a smaller domain, namely the half-plane $\Re x \geqq \alpha_0$ only, according to Theorem VI. Similar remarks will throw new light on the results of the author's recent paper, *Zur Abzählung der reellen Wurzeln algebraischer Gleichungen*, Math. Zeitschrift, vol. 38 (1934), pp. 546–564. The weaker upper bound is definitely worse as long as we restrict ourselves to real zeros. This is no longer true as soon as complex zeros are also taken into account, for in general weaker upper bounds hold for larger domains in the complex plane.

we get immediately, taking into consideration Theorem VI, the following corollary (see Laguerre, loc. cit., p. 157).

THEOREM VII. *The number of those zeros of a polynomial $f(x)$ of degree n which lie in the half-plane $\Re x \leqq a$ does not exceed the number of variations in the sum*

$$f(a) - \binom{n}{1}f(a+h) + \binom{n}{2}f(a+2h) - \cdots + (-1)^n f(a+nh) \qquad (h > 0),$$

and the difference between this upper bound and the actual number of zeros is even.

This last theorem can be applied with particular ease to polynomials defined by interpolation, i.e., when the values $f(a + \nu h)$ are given. It shows that if $f(a)$ is sufficiently large, we are sure to have no zeros with $\Re x \leqq a$.

Theorem VI may also be applied to locate the zeros of the derivative of a rational function $f(x)$ with only real zeros and poles. This is best shown by an example. Let the rational function be

$$f(x) = \frac{x^6(x-2)^3}{(x+1)^{16}(x-1)^2}. \tag{10}$$

Besides the obvious multiple zeros $x = 0$ and $x = 2$, $f'(x)$ will admit the same zeros as

$$F(x) = \frac{f'(x)}{f(x)} = -\frac{16}{x+1} + \frac{6}{x} - \frac{2}{x-1} + \frac{3}{x-2}.$$

Laguerre's extended theorems readily give the following results. There are no zeros with $\Re x \leqq -1$ and there is exactly one zero with $\Re x > 2$. Furthermore, there are no zeros within the circles of diameters $(-1, 0)$ and $(1, 2)$, while there are two zeros or none within the circle on $(0, 1)$. A direct solution is possible and gives the zeros

$$x = 3 \text{ and } x = (11 \pm i\sqrt{23})/18,$$

the complex roots being actually in the last mentioned circle. By a previous remark it becomes apparent that if $x = \alpha$ is the smallest pole or zero of a rational function $f(x)$ of the form (10), then $f'(x)$ will have no zeros with $\Re x \leqq \alpha$ $(x \neq \alpha)$, provided the order of the pole or zero $x = \alpha$ is sufficiently high.

8. It might be of interest to point out finally that Theorem VI can be extended by usual continuity considerations to infinite series and integrals of the type

$$\begin{aligned} F_1(x) &= \sum_{\nu=0}^{\infty} \frac{a_\nu}{x + p_\nu} \qquad (p_0 = 0 < p_1 < \cdots < p_\nu \to \infty), \\ F_2(x) &= \int_{-\infty}^{0} \frac{\varphi(t)\,dt}{x - t}, \end{aligned} \tag{11}$$

giving estimates for the numbers of zeros with $\Re x > 0$.

Concerning series of the type (11), the following results are readily derived from classical theorems of Abel. If the series converges in one point of the plane, it converges in the whole plane except the points $x = 0, -p_1, -p_2, \cdots$, and it converges uniformly in every closed and finite domain which is free of these points. The case of convergence occurs if and only if $\sum_0^\infty (a_\nu/p_\nu)$ converges. From these results and Theorem VI the following is readily derived.

THEOREM VIII. *The number of zeros of the sum $F(x)$ of the convergent series*

$$F(x) = \sum_{\nu=0}^{\infty} \frac{a_\nu}{x + p_\nu} \qquad (p_0 = 0 < p_1 < \cdots < p_\nu \to \infty)$$

within the half-plane $\Re x > 0$ does not exceed the number of variations in the infinite sum $a_0 + a_1 + a_2 + \cdots$.

If, for example, $\sum_0^\infty a_\nu = s \neq 0$, then $v(a_0 + a_1 + \cdots)$ is certainly finite and therefore $Z\{\Re x > 0\}$ is finite also. In the case $s \neq 0$, it is again seen by Abel's theorem that $F(x)$ has the sign of s for sufficiently large x. *Hence in this case* $(\sum_0^\infty a_\nu \neq 0)$ *the difference $v(a_0 + a_1 + \cdots) - Z\{\Re x > 0\}$ is always finite and an even number*, for $F(+0)$ has the sign of a_0, while $F(x)$ has the sign of $s = \sum_0^\infty a_\nu$ for large $x > 0$, as already remarked.

The two series[8]

$$\int_0^1 \frac{t^{x-1}\,dt}{t+1} = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{x+\nu}, \qquad \int_0^1 e^{-t} t^{x-1}\,dt = \sum_{\nu=0}^{\infty} \frac{1}{\nu!}\frac{(-1)^\nu}{x+\nu} \qquad (\Re x > 0)$$

are examples with $v(a_0 + a_1 + \cdots) = 0$.

§2. Proofs of Theorems II, III and IV extending the theorem of Descartes

9. A rotation of the a-plane about its origin does not change equation (1); hence we may assume that the imaginary axis bisects the separating double sector S of aperture ψ $(0 < \psi \leqq \pi)$. The remaining double sector (containing the points a_ν) is now bisected by the real axis, and its aperture is $2\eta = \pi - \psi$ $(\geqq 0)$. Hence $\psi/n = (\pi - 2\eta)/n$, and we have to prove the inequality

$$Z\{|\arg x| < (\pi - 2\eta)/n\} \leqq v_a(S). \tag{12}$$

The number of zeros of $f(x)$ inside the sector D: $|\arg x| < (\pi - 2\eta)/n$ is the same as for the function[9]

$$F(x) = e^{\frac{\pi i}{2}} x^{-\frac{n}{2}} f(x). \tag{13}$$

Let us decompose $F(x)$ into its real and imaginary parts using polar coördinates.

[8] The zeros of these two functions which are connected with the Gamma function were investigated by many authors and finally located by T. H. Gronwall, Trans. Amer. Math. Soc., vol. 28 (1926), pp. 391–399 and Annales de l'École Norm. Sup., vol. 33 (1916), pp. 381–393.

[9] Obreschkoff proved Theorem I by applying Cauchy's theorem directly to $f(x)$. In proving Theorem II by this method, the additional factor $x^{-n/2}$ seems to be essential.

Let

(14) $a_\nu = \rho_\nu e^{i\alpha_\nu}, \quad |\alpha_\nu| \leqq \eta, \quad \rho_\nu \gtreqless 0 \quad (\nu = 0, 1, \cdots, n), \quad x = re^{i\varphi} \quad (r \geqq 0).$

From (1) and (13) we get

$$\begin{aligned} F(x) &= \sum_{\nu=0}^{n} \rho_\nu e^{\left(\frac{\pi}{2} - \frac{n}{2}\varphi + \nu\varphi + \alpha_\nu\right)i} r^{-\frac{n}{2}+\nu} \\ &= \sum_{\nu=0}^{n} \rho_\nu \cos\left(\frac{\pi}{2} - \frac{n}{2}\varphi + \nu\varphi + \alpha_\nu\right) r^{-\frac{n}{2}+\nu} \\ &\quad + i\sum_{\nu=0}^{n} \rho_\nu \sin\left(\frac{\pi}{2} - \frac{n}{2}\varphi + \nu\varphi + \alpha_\nu\right) r^{-\frac{n}{2}+\nu} \\ &= P + iQ. \end{aligned} \tag{15}$$

There is obviously no restriction in assuming $a_0 a_n \neq 0$. *Let us furthermore assume for the moment that we have not only* $|\alpha_\nu| \leqq \eta$, *but also*

$$|\alpha_\nu| < \eta, \tag{16}$$

and that $f(x)$ *does not vanish on the two half-lines bounding sector* $D: |\arg x| < \psi/n$.

Draw about the origin two circles of radii R and ϵ (very large and very small respectively) which cut the boundary of D in the points A, B, A', B'. Consider the finite domain D' which is bounded by the straight segments $A'A$ and BB' and the circular arcs $\widehat{AB}$ and $\widehat{A'B'}$ and whose boundary is described counter-clockwise in the order $ABB'A'A$. We may assume $F(x) \neq 0$ along the boundary of D'.

Let x describe the boundary of D' counter-clockwise and consider the variation of the real function Q/P along this boundary. By a theorem of Cauchy we know that $2Z\{D'\}$ is equal to the number of zeros of Q/P in which Q/P passes from negative to positive values minus the number of its zeros in which this function passes from positive to negative values.[10]

Let us investigate the number of zeros of Q along the boundary of D'. Along the arc AB we have $-(\pi - 2\eta)/n \leqq \varphi \leqq (\pi - 2\eta)/n$; hence $\eta \leqq (n\varphi/2) + (\pi/2) \leqq \pi - \eta$, and finally

$$(0 <)\ \eta + \alpha_n \leqq \pi/2 + n\varphi/2 + \alpha_n \leqq \pi - \eta + \alpha_n\ (< \pi),$$

on account of (16). Within this range of values of φ $\cot(\pi/2 + n\varphi/2 + \alpha_n)$ is a finite and continuous function of φ, and (15) shows that

$$\lim_{r\to\infty} P/Q = \cot(\pi/2 + n\varphi/2 + \alpha_n)$$

[10] If $F(x)$ were a polynomial, Cauchy's theorem stated above would be a rather simple consequence of the fundamental theorem of algebra as shown by Ch. Sturm, Journal de Mathématiques, vol. 1 (1836), pp. 290–294. Our function (14) differs from a polynomial by the factor $x^{-n/2}$ only. The origin being outside of D', Sturm's elementary proof applies also to our case.

uniformly in φ within this range. Hence P/Q is finite and continuous along $\widehat{AB}$ provided R is sufficiently large. *In particular, Q will not vanish along AB.* Similarly

$$\lim_{r \to 0} P/Q = \cot (\pi/2 - n\varphi/2 + \alpha_0)$$

uniformly in φ for $-(\pi - 2\eta)/n \leqq \varphi \leqq (\pi - 2\eta)/n$. This shows that *$Q$ will not vanish along $\widehat{A'B'}$ if ϵ is sufficiently small.*

We want to show now that the sine coefficients of Q, namely $\sin (\pi/2 - n\varphi/2 + \nu\varphi + \alpha_\nu)$ are all positive along the boundary of our sector D, i.e., for $\varphi = \pm(\pi - 2\eta)/n$. For taking in each term of the sum $\frac{1}{2}\pi + (-\frac{1}{2}n + \nu)\varphi + \alpha_\nu$ respectively the smallest and the largest possible value we have

$$0 = \frac{\pi}{2} - \frac{n}{2}\frac{\pi - 2\eta}{n} - \eta < \frac{\pi}{2} + \left(-\frac{n}{2} + \nu\right)\varphi + \alpha_\nu < \frac{\pi}{2} + \frac{n}{2}\frac{\pi - 2\eta}{n} + \eta = \pi$$

for $\nu = 0, 1, \cdots, n$. By applying Descartes' rule to the polynomial in r

$$r^{n/2}Q = \sum_{\nu=0}^{n} \rho_\nu \sin\left(\frac{\pi}{2} - \frac{n\varphi}{2} + \nu\varphi + \alpha_\nu\right) r^\nu$$

for $\varphi = \pm(\pi - 2\eta)/n$, we see that *$Q$ can not vanish along the sides BB' and $A'A$ more than $2v_\rho = 2v_a(S)$ times.* By a proper choice of ϵ and R we may assume that $Z\{D'\} = Z\{D\}$, and combining our results with Cauchy's theorem we get

$$2Z\{D'\} = 2Z\{|\arg x| < (\pi - 2\eta)/n\} \leqq 2v_a(S).$$

This proves the inequality (12).

Our additional assumptions (16) are now easily removed by slightly increasing η (decreasing ψ). For now (16) certainly hold, and the increase of η (if sufficiently small) will remove possible zeros of $f(x)$ along the boundary of D without decreasing $Z\{D\}$. Now (12) is proved as before and η may be decreased to its original value.

In order to show that (3) may fail to hold if we add to the boundary of D the point $e^{-(\psi/n)i}$, say, consider the equation

$$e^{(\pi-\psi)i} + x^n = 0 \qquad (0 < \psi \leqq \pi). \tag{17}$$

The two lines through the origin containing the points $x = 1$ and $x = e^{(\pi-\psi)i}$ define a separating double sector S of aperture ψ with $v_a(S) = 0$. Hence $Z\{|\arg x| < \psi/n\} = 0$. This relation will become false if we add the point $x = e^{-(\psi/n)i}$ to D, for this point is a zero of the equation (17).[11]

[11] It should be remarked that D ceases to be the largest domain for which (3) holds if, besides ψ and n, $v_a(S)$ also has a prescribed value > 0. Thus for real equations ($\psi = \pi$) Obreschkoff really proved more than $Z\{|\arg x| < \pi/n\} \leqq v_a$, namely, $Z\{|\arg x| < \pi/(n - v_a)\} \leqq v_a$, in his paper *Über die Wurzeln algebraischer Gleichungen*, Jahresbericht der Deutschen Math.-Vereinigung, vol. 33 (1924), p. 61. It would be interesting to improve the inequality (3) accordingly.

10. In proving Theorem III there is obviously no loss of generality in assuming all the coefficients $a_\nu \neq 0$. Take a regular pencil Σ of $n + 1$ straight lines through 0 which divide the plane in $2n + 2$ equal angles of aperture $\pi/(n + 1)$. Let one of the lines of the pencil pass through a_0. The pencil defines $n + 1$ double sectors (each composed of a pair of opposite sectors) and the n segments $0a_1, 0a_2, \cdots, 0a_n$ can obviously occupy the interiors of at most n of these double sectors. Hence at least one of them will be a separating double sector of aperture $\pi/(n + 1)$ for the points a_ν. Note on the other hand that the points $a_\nu = e^{\nu\pi i/(n+1)}$ $(\nu = 0, 1, \cdots, n)$ do admit separating double sectors of aperture $\pi/(n + 1)$ but none of greater aperture. Thus Theorem III is proved.

11. The proof of Theorem IV resembles that of Theorem II. It suffices to remark that Cauchy's theorem is applied to the function

$$\Phi(x) = e^{\frac{\pi}{2}i - \frac{\lambda_n - \lambda_0}{2}x} F(x)$$

along the boundary of the rectangular domain $|\Re x| \leqq A$ (A very large), $|\Im x| \leqq \psi/(\lambda_n - \lambda_0)$.

§3. Proof of Theorem VI extending the theorem of Laguerre

12. Let us first consider the case when $\sum_0^n a_\nu \neq 0$. We may assume $\alpha_0 = 0$, and hence $\alpha_n < \alpha_{n-1} < \cdots < \alpha_1 < \alpha_0 = 0$. For the rational function $F(x)$ defined by (7) we have to prove the inequality

$$Z\{\Re x \geqq 0\} \leqq v(a_0 + a_1 + \cdots + a_n). \tag{18}$$

Take $\epsilon > 0$ so small that $F(x)$ does not vanish on the line $\Re x = -\epsilon$ and $Z\{\Re x > -\epsilon\} = Z\{\Re x \geqq 0\}$. Draw two circles about the origin of radii 2ϵ and R (very large) which intersect the line $\Re x = -\epsilon$ in the points A', B' and A, B, respectively, and consider the domain D': $\Re x \geqq -\epsilon$, $2\epsilon \leqq |x| \leqq R$, whose boundary is described counter-clockwise in the order $ABB'A'A$. For ϵ sufficiently small and R sufficiently large we have $Z\{D'\} = Z\{\Re x \geqq 0\}$.

Let

$$x = re^{i\varphi}, \qquad x - \alpha_\nu = r_\nu e^{i\varphi_\nu}, \qquad p_\nu = -\alpha_\nu = |\alpha_\nu|. \tag{19}$$

From (7) and (19) we get

$$F(x) = \sum_{\nu=0}^{n} \frac{a_\nu}{r_\nu} e^{-i\varphi_\nu} = \sum_{\nu=0}^{n} a_\nu \frac{\cos\varphi_\nu}{r_\nu} - i \sum_{\nu=0}^{n} a_\nu \frac{\sin\varphi_\nu}{r_\nu} = P + iQ. \tag{20}$$

We wish to apply Cauchy's theorem to $F(x)$ in D', and for this purpose let us investigate the zeros of Q/P along the boundary $ABB'A'A$ of D'.

(i) If R is sufficiently large, along the circular arc $\overset{\frown}{AB}$ Q/P will have the sign of $-\tan\varphi$ and will therefore vanish only once, for $\varphi = 0$, and pass there from positive to negative values.

(ii) Similarly along $\widehat{B'A'}$ Q/P will have, if ϵ is sufficiently small, the sign of $-\tan \varphi$ and will therefore vanish once, for $\varphi = 0$, and pass there from negative to positive values.

(iii) Let us now move the point x along BB'. If we put $x = -\epsilon + it$, from the triangle $(\alpha_\nu, -\epsilon, x)$ we get $\sin \varphi_\nu = t/r_\nu$ or $\sin \varphi_\nu/r_\nu = t/r_\nu^2 = t/[t^2 + (p_\nu - \epsilon)^2]$. Hence

$$Q = -t\left(\frac{a_0}{t^2 + \epsilon^2} + \frac{a_1}{t^2 + (p_1 - \epsilon)^2} + \cdots + \frac{a_n}{t^2 + (p_n - \epsilon)^2}\right).$$

Hence if $\epsilon^2 < (p_1 - \epsilon)^2$, as we may assume, and if we take t^2 as the new variable, Laguerre's Theorem V shows that Q has at most $v(a_0 + a_1 + \cdots + a_n)$ zeros along the half-line $x = -\epsilon + it$ $(t > 0)$, which implies that Q/P has at most $v(a_0 + \cdots + a_n)$ zeros along BB'.

(iv) The above result holds along $A'A$ as well because $Q(-\epsilon - it) = -Q(-\epsilon + it)$.

By Cauchy's theorem, therefore, we have

$$2Z\{\Re x \geqq 0\} = 2Z\{D'\} \leqq 2v(a_0 + \cdots + a_n) + 1 - 1 = 2v(a_0 + \cdots + a_n).$$

This proves (18).

Assuming now that $\sum_0^n a_\nu = 0$, let us prove that

$$Z\{\Re x > 0\} \leqq v(a_0 + \cdots + a_n). \tag{21}$$

Let $\sum_0^{n-1} a_\nu$ be positive. A sufficiently slight increase of a_n will not change $Z\{\Re x > 0\}$. Now $\sum_0^n a_\nu > 0$ and (18) holds. Hence (21) holds à fortiori.

In conclusion, let me point out that *if* $\alpha_0 = 0$ *and all the other constants occurring in* (7) *are variable (with* $\sum_0^n a \neq 0$), *then* $\Re x \geqq 0$ *is the largest domain for which the inequality* (8) *holds.* This inequality may become false if we add to $\Re x \geqq 0$ even a single point with $\Re x < 0$. This is best shown as follows. With $p_\nu = -\alpha_\nu$, $b_\nu = (p_{\nu+1} - p_\nu) \sum_0^\nu a_k$ $(p_{n+1} = p_n + 1)$, we have

$$F(x) = \frac{b_0}{x(x + p_1)} + \frac{b_1}{(x + p_1)(x + p_2)} + \cdots + \frac{b_{n-1}}{(x + p_{n-1})(x + p_n)} + \frac{b_n}{x + p_n}, \tag{22}$$

$$(0 < p_1 < \cdots < p_n).$$

If all the b_ν are $\geqq 0$, we know from (8) that $Z\{\Re x \geqq 0\} = 0$. By a certain type of argument used by J. v. S. Nagy and Morris Marden[12] it is readily shown that if $b_\nu \geqq 0$ and all the constants occurring in (22) are variable, then $\Re x < 0$ is the geometric locus of the zeros of $F(x)$. This proves our last remark.

INSTITUTE FOR ADVANCED STUDY.

[12] J. v. S. Nagy, *Über die Lage der Wurzeln von linearen Verknüpfungen algebraischer Gleichungen*, Acta Szeged, vol. 1 (1923), pp. 127–138, and M. Marden, loc. cit.

CONNECTIONS BETWEEN DIFFERENTIAL GEOMETRY AND TOPOLOGY

II. CLOSED SURFACES

By Sumner Byron Myers

1. Introduction. This paper deals with closed 2-dimensional Riemannian manifolds, for brevity designated *closed surfaces*. The properties of a fundamental locus which we call the minimum point locus with respect to a point A, studied in a previous paper[1] by the author for the case of simply-connected analytic surfaces, are determined here for the general class of closed surfaces. The locus in question is defined as the locus m of points M on geodesic rays issuing from a point A, which are the last points along these rays such that the arc AM furnishes an absolute minimum (proper or improper) to the length of arcs joining A to M. In the case of a closed analytic surface S the principal result is that m is a linear graph (i.e., a finite connected 1-dimensional complex) whose one-dimensional Betti number equals the one-dimensional Betti number mod 2 of the surface. A study is made of the parametrization of m by means of θ, the angular coördinate of the geodesic rays through A. It is found that this depends on the orientability of S, and also that the number of values of θ yielding one point of m equals the order[2] of that point in m.

A brief study is also made of non-analytic surfaces. Here we assume, for example, that S is a closed regular manifold of[3] class 5 with a Riemannian line element of class C^4. The locus m turns out to be a continuous curve (not necessarily a linear graph) with the same relation as in the analytic case between the one-dimensional Betti number of m and the one-dimensional Betti number of S, and similar relations among the orientability of S, the parametrization of m by means of θ, and the order of points of m.

In both analytic and non-analytic cases, if we subtract the locus m from the surface S, the result is a single 2-cell with m as its singular boundary, simply covered (except at A) by the geodesic rays through A. Thus is solved the prob-

Received October 4, 1935; presented to the American Mathematical Society April 19, 1935. The author is National Research Fellow.

[1] See Myers, *Connections between differential geometry and topology, I. Simply connected surfaces*, this Journal, vol. 1 (1935), pp. 376–391. This paper will be referred to as (I). An abstract containing the results of that paper, as well as the results of the present paper in the analytic case, appears in the Proc. Nat. Acad. Sci., April, 1935, under the same title.

[2] By the order of a point P of a continuous curve C we mean the number of 1-cells contained in C which issue from P and are such that no two of them have any point in common but P.

[3] See Veblen and Whitehead, *The Foundations of Differential Geometry*, p. 81.

lem of finding the largest domain of the geodesic polar coördinates and normal coördinates with A as pole.

The locus in question was originally introduced by Poincaré,[4] who considered only closed simply connected surfaces of positive curvature. Poincaré called the locus "lignes de partage". J. H. C. Whitehead[5] considers the same locus in a recent paper. He calls it the "cut locus", and considers it on complete n-dimensional Finsler spaces. He obtains the theorem that such a space can be decomposed into an n-cell and the cut locus, which forms the singular boundary of the n-cell, but he does not study the topological nature of the locus itself. The results that we obtain in the present paper for 2-dimensional Riemannian spaces can readily be extended to 2-dimensional Finsler spaces. It is probable that analogous equalities between the Betti numbers of the space itself and the Betti numbers of the locus exist in n dimensions. However, it seems difficult to prove that the locus in the analytic n-dimensional case is homeomorphic to a finite $(n - 1)$-dimensional complex.

2. The analytic case. We assume a knowledge of the definitions and terminology of (I). We recall that on a closed (compact) surface every pair of points can be joined by a geodesic of class $\mathfrak{A}$, i.e., a geodesic every segment of which furnishes an absolute minimum, proper or improper, to the length of arcs joining its end points. Lemmas 1–11 and Theorems 1–3 of (I) hold here as well as in the simply connected case.

Theorem 4 of (I) is replaced by the following

THEOREM 1. *A surface is closed if and only if there exists no geodesic ray of class $\mathfrak{A}$ through any point on it.*

For Rinow has proved[6] that from every point of an open surface issues a geodesic ray of class $\mathfrak{A}$. Furthermore, on a closed surface there can exist no infinite set of points without limit point, and on a geodesic ray of class $\mathfrak{A}$ such a set exists.

Thus on every geodesic ray through A on S there is a minimum point with respect to A. It follows from Lemma 8 of (I) that the locus m of minimum points with respect to A is the continuous single-valued image of a circle; i.e., a continuous curve.

Now if we cut off each geodesic ray from A at its minimum point with respect to A, the region $S - m - A$ is simply covered by these truncated geodesic rays. For if two such geodesic rays intersected before m, the absolute minimum property would have stopped on at least one of them at or before the intersection with the other, as follows from Lemma 4 of (I). Hence we have

THEOREM 2. *If the minimum point locus m with respect to A is deleted from S, the result $S - m$ is a 2-cell σ with m as its singular boundary. σ is simply covered*

[4] Trans. Amer. Math. Soc., vol. 6 (1905), p. 243.

[5] Annals of Math., vol. 36 (1935), pp. 679–704.

[6] W. Rinow, *Ueber Zusammenhänge zwischen der Differentialgeometrie im Grossen und im Kleinen*, Math. Zeitschrift, vol. 35 (1932), p. 522.

(except at A) by the geodesic rays issuing from A, and hence forms the largest domain of the geodesic polar coördinates and normal coördinates with A as pole.

No set C of closed curves on m can bound on S. For C would separate S into two regions R_1 and R_2 such that no point in R_1 could be joined to a point in R_2 without crossing C. But A would be in either R_1 or R_2, say R_1, and could be joined to any point in R_2 by a geodesic of class $\mathfrak{A}$. This geodesic would cross C, thus contradicting the fact that it is of class $\mathfrak{A}$. Let B_1 be the one-dimensional Betti number mod 2 of S. Then we have shown that the one-dimensional Betti number of m, i.e., the number of independent sets of closed curves in m, is at most B_1. On the other hand, the one-dimensional Betti number of m is at least B_1. Otherwise a non-bounding closed curve independent with respect to homology of those of m could be drawn on S. This would contradict Theorem 2. Thus the one-dimensional Betti number of m is exactly equal to B_1. Hence

THEOREM 3. *The minimum point locus m with respect to any point A of a closed surface S is a continuous curve whose one-dimensional Betti number equals the one-dimensional Betti number mod 2 of S.*

According to a theorem of curve theory, m is locally a tree[7] because its one-dimensional Betti number is finite. We shall prove now that in the analytic case the number of end points (i.e., points of order 1) of m is finite. From this it will follow that m is a linear graph,[8] i.e., a finite connected 1-dimensional complex.

By a minimum point of order n with respect to A we mean a point of m which can be joined to A by just n geodesics of class $\mathfrak{A}$. We prove first that every end point P of m is a minimum point of order 1.

Suppose A could be joined to P by two geodesics of class $\mathfrak{A}$, g_1: $\theta = \theta_1$ and g_2: $\theta = \theta_2$. We could find a neighborhood σ of P of radius δ which would be divided into two 2-cells σ_1 and σ_2 by these two geodesic rays. By Lemma 8 of (I) we can find an ϵ so small that the geodesic rays for $\theta_1 - \epsilon < \theta < \theta_1 + \epsilon$ all have their minimum points with respect to A within the neighborhood σ. The rays for $\theta_1 - \epsilon < \theta < \theta_1$ (if ϵ is small enough) all remain close to g_1, by Lemma 1 of (I), and all enter one of the regions σ_1 or σ_2, say σ_1. The minimum points on these geodesic rays are all in σ_1, for none of these rays can cross g_1 or g_2 and remain of class $\mathfrak{A}$. Similarly, geodesic rays from A whose θ-coördinates lie between θ_1 and $\theta_1 + \epsilon$ have their minimum points in σ_2. Thus the locus m has two distinct continuous curves issuing from P, or the rays for $\theta_1 - \epsilon < \theta < \theta_1 + \epsilon$ all have their minimum points with respect to A at P, in which case the complete locus m is the single point P, as shown in (I) on p. 387. Either of these cases contradicts the assumption that P is an end point of m. Hence every end point of m is a minimum point of order 1.

But there is only a finite number of minimum points of order 1 with respect to

[7] See Menger, *Kurventheorie*, p. 323.

[8] This follows from Menger, loc. cit., p. 266. It is easily shown that if a continuous curve has a finite number of end points and a finite number of closed curves, it has only a finite number of points of order greater than 2.

A on S. This is proved as in the simply connected case[9] in (I) by means of Lemmas 9 and 10 and Theorems 2 and 3 of (I). Hence m has only a finite number of end points, and m is a linear graph. This linear graph is already partly triangulated by means of the end points and branch points, but it may be necessary to subdivide the closed curves of m in order to make m a complex in the technical sense.

From Lemma 11 of (I) it follows that any arc of m containing no points conjugate to A and no interior points of order > 2 is a regular analytic arc.

THEOREM 4. *The locus m on a closed analytic surface is a linear graph. The end points of m are conjugate to A, and are cusps turned toward A of the locus of first points conjugate to A. An arc of m containing no points conjugate to A and no interior points of order > 2 is a regular analytic arc.*

In the proof of Theorem 4, we have shown that every end point of m is a minimum point of order 1. We now show, conversely, that every minimum point of order 1 is an end point, thus identifying the points of m of order 1 with the minimum points of order 1. More generally, we will show that the order of a point of m equals its order as a minimum point with respect to A.

Let $\bar{P}$ be an arbitrary minimum point of order 1 with respect to A on the geodesic ray $\theta = \bar{\theta}$. Draw a geodesic circle γ of radius δ about $\bar{P}$ so small that it and its interior σ are simply covered by the geodesic rays from $\bar{P}$. If ϵ is chosen small enough, all the geodesic rays from A for $\bar{\theta} - \epsilon < \theta < \bar{\theta} + \epsilon$ remain close to $\theta = \bar{\theta}$ and by Lemma 8 have their minimum points with respect to A in σ. Hence none of them have points conjugate to A before or when they reach γ, so that they form a field F in the neighborhood of $\theta = \bar{\theta}$ up to and including a portion of γ. It is easily seen that if two geodesic rays from A in the interval $\bar{\theta} - \epsilon < \theta < \bar{\theta} + \epsilon$ intersect again in σ, they bound a 2-cell lying in $F + \sigma$.

Since the number of minimum points of order 1 is finite, there exists an interval $\theta_1\theta_2$ containing $\bar{\theta}$ but no other value of θ which furnishes a minimum point of order 1 with respect to A. Each value of θ except $\bar{\theta}$ in this interval furnishes a minimum point of order > 1, and hence to each such value of θ corresponds another value of θ, say θ', which furnishes the same minimum point P. If θ is close enough to $\bar{\theta}$, then P is very close to $\bar{P}$, by Lemma 8 of (I), and from Lemma 5 and the fact that P is of order 1 we conclude that θ' is very close to $\bar{\theta}$. Thus if θ is close enough to $\bar{\theta}$, both θ and θ' lie in $\theta_1\theta_2$ and also in the interval $\bar{\theta} - \epsilon < \theta < \bar{\theta} + \epsilon$. But according to the conclusion of the previous paragraph, the rays θ and θ' bound a 2-cell on S, and the same reasoning as that used in (I), p. 388, shows that the interval $\theta\theta'$ must include a value of θ furnishing a minimum point of order 1. Hence $\theta_1 < \theta < \bar{\theta} < \theta' < \theta_2$ or $\theta_1 < \theta' < \bar{\theta} < \theta < \theta_2$. But in (I), pp. 388–389, it was proved that if the rays θ and θ' bound a 2-cell containing just one minimum point P of order 1, that point P is an end point of m. Thus every minimum point of order 1 is an end point of m, and the set of minimum points of order 1 with respect to A is identical with the set of end points of m.

We now prove by induction that the order of a point of m equals its order as a

[9] See (I), p. 387 (top) and p. 388 (bottom).

minimum point with respect to A. Assuming the proposition true for all integers $n < q$, we shall prove it for $n = q$. A point P of m of order q cannot be a minimum point of order $w > q$; for it can be shown that from the latter type of point issue w distinct 1-cells of the linear graph m in the same way that we showed that from a minimum point of order 2 issue two distinct 1-cells of m. Furthermore, P cannot be a minimum point of order $< q$, for part of our induction hypothesis is that a minimum point of order $z < q$ is a point of order z of m. Thus the order of a point of m equals its order as a minimum point with respect to A, and hence we have the following

THEOREM 5. *The order of a point of m equals its order as a minimum point with respect to A. In particular, the end points of m are identical with the minimum points of order* 1 *with respect to A.*

On the basis of this theorem, we see that if we parametrize the continuous curve m in terms of θ, the order of a point of m must equal the number of values of θ furnishing that point. As θ ranges from 0 to 2π each 1-cell of m is traced out twice, while each 0-cell of m is covered a number of times equal to its order.

The linear graph m is the singular image of a simple closed curve $\gamma: r = r(\theta)$ around the pole in the euclidean (r, θ)-plane. The whole surface S can be got topologically by considering γ and its interior with identification of certain pairs of 1-cells on γ. But it is well known[10] that if an orientation is given to a polygon, the manifold constructed by considering the polygon and its interior with identification of pairs of sides of the polygon is non-orientable or orientable according as to whether or not the identification of at least one pair of sides of the polygon is made with the same orientation of the two sides concerned. Hence we have

THEOREM 6. *If S is orientable, as θ increases from* 0 *to* 2π *every* 1*-cell of m is traced out twice, once in each sense. If S is non-orientable, at least one* 1*-cell is traced out twice in the same sense.*

Thus the connectivity and orientability of the 2-dimensional manifold S, and hence the complete topology of S, can be determined from a knowledge of the 1-dimensional Betti number of m and the way in which m is traced out. Another way of stating this follows. The number of independent closed curves in m determines the number of generators in the fundamental group G of S, while the manner in which m is traced out determines the generating relation of G. Conversely, a knowledge of the topology of S determines the one-dimensional Betti number of m and to some extent the way in which m is traced out.

3. The non-analytic case. We now consider briefly the case where the surface S satisfies certain differentiability conditions, but is not necessarily analytic. Suppose that S is a closed regular 2-dimensional manifold of class 5 with a line element of class C^4.

Lemma 1 of (I) holds here with the change that $x(r, \theta)$ and $y(r, \theta)$ are no longer analytic, but only functions of class C^3. In Theorem 1 of (I) the function $f(r, \theta)$ continued in the same way as in the analytic case becomes a function of r, θ of

[10] See, for example, Seifert and Threlfall, *Topologie*, p. 135.

class C^2 for all θ and all $r > 0$. This is also true of the function $K(r, \theta)$. As for Theorems 2 and 3 of (I), the locus of first conjugate points to A is again a single point, a closed curve, or a set of one or more open curves, parametrizable as functions of class C^2 of θ. There may now be an infinite number of cusps of the locus even on a finite segment of it, but the number of curves in the locus can still be infinite only if a curve of the set can be found all of whose points are arbitrarily far from A on the geodesics on which they are conjugate to A. Since the values of θ furnishing singular points of the locus form a closed set, such values cannot be dense on any interval I in $0 \leqq \theta \leqq 2\pi$ unless all the values of θ in I furnish the same conjugate point, as seen from (3.10) of (I). Lemmas 4–9 of (I) hold unchanged. Lemma 10 must be modified to allow the possibility that M is a limit point of cusps of the locus of first conjugate points to A. In Lemma 11, the arc d now becomes a regular arc of class C^3.

Theorems 1, 2, and 3 of the present paper in no way depend on analyticity. Hence they hold for the more general class of surfaces. Thus the minimum point locus with respect to a point A is a continuous curve whose one-dimensional Betti number equals the connectivity number mod 2 of S. Hence it is locally a tree, and therefore a regular curve in the sense of curve theory.[11] Theorem 4, however, does not hold in the non-analytic case, for the proof of the finiteness of the number of end points, which has as a consequence that m is a linear graph, depends essentially on the analyticity of S. We can, however, give some restrictions on the number of end points of m.

In the first place, the method used in the proof of Theorem 4 can be used here to show that if A can be joined to P by two geodesics of class $\mathfrak{A}$, then either two distinct 1-cells of m issue from P or else a whole interval of values of θ furnish geodesic rays of class $\mathfrak{A}$ joining A to P. In the latter case we shall say that P is a minimum point of order I. Thus every end point P of m is either a minimum point of order 1 or a minimum point of order I. In the first case, P is conjugate to A on the unique geodesic of class $\mathfrak{A}$ joining it to A, by Lemma 9 of (I), while in the second case P is conjugate to A on each geodesic ray of the whole closed interval of rays of class $\mathfrak{A}$ joining A to P. Furthermore, such a point P is always a singular point of the conjugate point locus C; in fact, it is either a cusp of C or a limit point of cusps of C. From the reasoning used a few paragraphs back, the set of values of θ furnishing end points of m cannot be everywhere dense in $0 \leqq \theta < 2\pi$, nor can it be dense on any subinterval unless all values of θ on that subinterval furnish just one end point.

According to our modification of Lemma 11 of (I), an arc of m containing no point conjugate to A and no interior point of order > 2 is a regular arc of class C^3.

Theorem 5 also is not exactly true for the non-analytic case. We have already seen that every end point of m is either a minimum point of order 1 or of order I. A minimum point of order I may have any order whatever as a point of m. However, we can still prove that a minimum point of order 1 is an end point of m. In the first place, this is true for an isolated minimum point of order 1 by

[11] See Menger, loc. cit., p. 96.

exactly the same proof as in the analytic case. For a general minimum point of order 1, we can give the following proof.

Let $\bar{P}$ be the minimum point of order 1, and suppose that $\bar{P}$ is of order > 1 as a point of m. Then from $\bar{P}$ issue at least two distinct arcs e_1 and e_2 contained in m. If the neighborhood σ of $\bar{P}$ used in the proof of Theorem 5 is small enough, e_1 and e_2 together divide σ into two parts σ_1 and σ_2. One of these, say σ_1, contains part of the geodesic ray $\theta = \bar{\theta}$ on which $\bar{P}$ is a minimum point with respect to A. For $\bar{\theta} - \epsilon < \theta < \bar{\theta} + \epsilon$, the minimum points with respect to A lie in σ. Furthermore, since the geodesic rays from A in this interval remain close to $\theta = \bar{\theta}$, all these minimum points lie in σ_1 or on $e_1 + e_2$, for the rays cannot cross $e_1 + e_2$ without losing their minimum property. Thus any geodesic arc from A for $\bar{\theta} - \epsilon < \theta < \bar{\theta} + \epsilon$ joining A to a point in σ_2 cannot be of class $\mathfrak{A}$. Consider a sequence of points P_i in σ_2 approaching $\bar{P}$, and join them to A by geodesics $\theta = \theta_i$ of class $\mathfrak{A}$. By Lemma 6 of (I), $\theta_i \to \bar{\theta}$. This gives a contradiction, and hence $\bar{P}$ cannot be of order > 1 as a point of m. Hence P is an end point of m.

Following the same induction method used in the analytic case, we can prove that a point of order n (n finite) of the locus m is either a minimum point of order n with respect to A or a minimum point of order $n - q + qI$, where $q \leqq n$; in other words, a point of finite order n of m has as its inverse image n connected pieces of the θ axis. If P is a point of[12] order ω of m, it is a minimum point of order $M + NI$, where $M + N \geqq \aleph_0$. P cannot be a point of order $> \omega$ of m, for m is a regular curve.

As for Theorem 6, by a rather complicated proof it can be shown that the following is true in the non-analytic case. As θ increases from 0 to 2π, if S is orientable, every 1-cell contained[13] in m which has no cyclic branch[14] is traced out twice, once in each sense. If S is non-orientable, at least one 1-cell in m is traced out twice in the same sense.

4. **Examples.** As a first example, let us consider the projective plane of constant positive curvature. This manifold is obtained from a sphere by identifying diametrically opposite points. Let A be any point on a sphere S of radius a, A' the opposite pole. Then since A and A' become identical on the projective plane p obtained from S, every geodesic through A on p is closed, and of length πa. The minimum point with respect to A on a geodesic ray g issuing from A on p is at a distance of $\pi a/2$ from A along g. Thus the minimum point locus with respect to A on p consists of a circle, traced twice in the same sense as θ increases from 0 to 2π. The only point conjugate to A on p is A itself.

On the ordinary torus in 3-dimensional euclidean space the situation is more complicated.[15] Let A be a point on the outer equator. The minimum point

[12] See Menger, loc. cit., p. 100.

[13] By a 1-cell contained in m we mean a subset of m homeomorphic to a 1-cell.

[14] A 1-cell e contained in m is said to have no cyclic branch if every closed curve in m containing any point of e contains the whole of e.

[15] For a study of the geodesics on a torus, see Bliss, Annals of Mathematics, vol. 4 (1902–3), pp. 1–21.

locus with respect to A consists of (a) the inner equator, (b) the meridian circle through the point A' diametrically opposite to A on the outer equator, and (c) two arcs of the outer equator issuing in opposite directions from A'. Roughly speaking, the locus consists of two closed curves intersecting in one point plus two branches issuing from a point of one of the closed curves. The minimum point locus with respect to any point whatever of the torus contains two non-bounding, non-homologous closed curves; the locus with respect to any point of the inner equator consists entirely of two such closed curves, since the points on the inner equator are perfect poles (i.e., points without conjugate point).

On any orientable surface of genus $p > 0$, the locus with respect to any point contains $2p$ closed curves, which form a basis for the 1-dimensional homology group mod 2 of S as well as a set of generators for the fundamental group of S. With respect to a perfect pole of such a surface, the locus consists entirely of $2p$ closed curves. Thus on a closed orientable surface of zero curvature or constant negative curvature, where every point is a perfect pole, the locus with respect to any point whatever consists entirely of $2p$ closed curves. This has an obvious relation to the well-known method of representation of closed orientable surfaces as polygons of $4p$ sides and their interiors in the euclidean ($p = 1$) or hyperbolic ($p > 1$) planes with identification of pairs of sides in the proper manner.

PRINCETON UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY.

CONVEX POLYHEDRA AND CRITERIA FOR IRREDUCIBILITY

By Casper Shanok

1. **Introduction.** This paper gives an application of Minkowski's[1] theory of convex polyhedra to the construction of irreducibility criteria for polynomials in several variables, thus generalizing the results of Dumas[2] for polynomials in one variable obtained by the use of convex polygons. The results of Minkowski referred to concern the nature of the least convex polyhedron determined by the set of points $\{p\}$ where each point p is of the form $\sum s_i p_i$, each s_i a constant > 0 and each p_i a point of a given polyhedron K_i. As a special case, these results include the case of the polyhedron K determined by two given polyhedra K_1 and K_2 with $s_1 = s_2 = 1$. This is the case which is of interest to us for the results that follow. For reasons obvious later, we shall term K the product, rather than the sum, of K_1 and K_2 and shall denote this relation by the notation $K = K_1 \cdot K_2$.

2. **Decomposability.** If now we consider the converse problem, namely, given K to determine K_1 and K_2 such that $K_1 \cdot K_2 = K$, we find first of all that we must assume we are dealing with polyhedra whose vertices have integral coördinates, for otherwise K_1 and K_2 can be chosen in an infinity of ways, i.e., the problem has no meaning. We likewise impose the further restriction that neither of the factors of K shall be a point, since this would simply amount to a translation of K with no accompanying change in its shape. If, under these conditions, it is possible to determine two polyhedra K_1 and K_2 (either or both may be lines or polygons) such that $K_1 \cdot K_2 = K$, we say that K is decomposable. We proceed to set up necessary conditions for the decomposability of K by considering its projections on the coördinate planes.

First projecting the vertices of K on the xy-plane, we determine the least convex polygon containing this set of points. This polygon we name the xy-boundary polygon of K and denote by b_{xy}. We have at once the result that for K to decompose it is necessary that each of the three boundary polygons of K decompose.[3] For if K decomposes into K_1 and K_2, the product of the xy- (or

Received by the Editors of the Annals of Mathematics July 2, 1934, accepted by them, and later transferred to this journal; presented to the American Mathematical Society April 14, 1933. The author wishes to express his gratitude to Prof. O. Ore for his assistance in the preparation of this paper, which is an abstract of a dissertation presented for the degree of Doctor of Philosophy in Yale University.

[1] H. Minkowski, *Gesammelte Abhandlungen*, vol. 2, pp. 131–229.

[2] G. Dumas, *Sur quelques cas d'irréductibilité des polynômes à coefficients rationnels*, Journal de Mathématiques Pures et Appliquées, (6), vol. 2 (1906), pp. 191–258.

[3] By extending Dumas' result in the plane to cover the entire least convex polygon, it is possible to show that the decomposition of a plane polygon amounts to the division of the sides of the given polygon into two sets such that the sides of each set, translated, wherever necessary, form a closed convex polygon, each side keeping its outer normal unchanged.

xz or yz) boundary polygons of K_1 and K_2 is the xy- (or xz or yz) boundary polygon of K.

The condition stated above is, in fact, even more stringent, for not only is it necessary that b_{xy}, for example, decompose, but further that to each point α of the decomposition set[4] in question there correspond a lattice point on the side or sides of K whose projection contains α.

To obtain a stronger necessary condition, we next project the faces of the lower (or of the upper) part of K on the xy-plane, getting a network N of non-overlapping convex polygons filling up the interior of b_{xy}. Now assuming that K decomposes into the product of K_1 and K_2 and forming the two similar networks N_1 and N_2 for K_1 and K_2, respectively, let us determine the relation between these networks. In the first place, as stated above, the product of the xy-boundary polygons of K_1 and K_2 is identical with that of K. In the second place, it follows from Minkowski's work[5] that every elementary polygon[6] of N is either an elementary polygon of K_1 or K_2 translated, or the product of parts of two elementary polygons, one of K_1 and the other of K_2. This leads to the second necessary condition for the decomposability of K, namely, that it must be possible to decompose N (and each of the other five similarly determined networks) by decomposing each of its elementary polygons and piecing the decompositions together to form two similar networks.

In concluding our discussion of decomposability, mention should be made of the exceptional, though somewhat trivial, case that K has as a factor a line segment parallel to one of the axes, say the z-axis. Due to the fact that such a factor has a point as its projection on the xy-plane, the fact that the xy-boundary polygon (or that either of the two networks bounded by the xy-boundary polygon) is indecomposable (i.e., is indecomposable except for a possible point factor) need not imply that K is indecomposable, but only that K is indecomposable except for a possible line factor parallel to the z-axis.

3. **The general criterion for the case of three variables.**

Now applying the results above to the construction of criteria for irreducibility, let us first consider the case of polynomials of three variables with coefficients belonging to some fixed field. Letting $F(x, y, z) = \sum A_{\alpha\beta\gamma} x^\alpha y^\beta z^\gamma$ be such a polynomial, we plot the set of points $\{(\alpha, \beta, \gamma)\}$ and determine the least convex polyhedron containing these points. Regarding this polyhedron, which we shall term the polyhedron of F, we proceed to prove the following

THEOREM 1. *Let the reducible polynomial*

$$F(x, y, z) = F_1(x, y, z) \cdot F_2(x, y, z).$$

Then $K = K_1 \cdot K_2$, *where* K, K_1 *and* K_2 *are the polyhedra of* F, F_1 *and* F_2 *respectively.*

[4] By a decomposition set we mean the set of lattice points containing (1) all the vertices of b_{xy} and (2) such lattice points as mark the points of division of those sides of b_{xy} which are broken up in the decomposition of b_{xy} in question.

[5] Loc. cit., p. 186.

[6] By an elementary polygon we mean the projection of a face of K.

Proof. Let $F_1(x, y, z) = \sum A'_{\alpha'\beta'\gamma'} x^{\alpha'} y^{\beta'} z^{\gamma'}$

and $F_2(x, y, z) = \sum A''_{\alpha''\beta''\gamma''} x^{\alpha''} y^{\beta''} z^{\gamma''}$.

Then

$$F(x, y, z) = \sum_{\alpha',\beta',\gamma'} \sum_{\alpha'',\beta'',\gamma''} A'_{\alpha'\beta'\gamma'} A''_{\alpha''\beta''\gamma''} x^{\alpha'+\alpha''} y^{\beta'+\beta''} z^{\gamma'+\gamma''}. \tag{1}$$

In the expanded product (1) let there be m terms[7] containing $x^\alpha y^\beta z^\gamma$. Setting S equal to the sum of these terms, we have

$$S = (b_1 + \cdots + b_m)x^\alpha y^\beta z^\gamma = A_{\alpha\beta\gamma} x^\alpha y^\beta z^\gamma. \tag{2}$$

We note here that S may drop out if $m \geqq 2$, but not if $m = 1$.

A study of the above relations shows that proving $K = K_1 \cdot K_2$ amounts to proving that the sets $\{(\alpha, \beta, \gamma)\}$ and $\{(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')\}$ determine the same polyhedron. From relations (1) and (2), it follows at once that every member of the set $\{(\alpha, \beta, \gamma)\}$ is a member of the set $\{(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')\}$, i.e., K is actually contained in or equal to $K_1 \cdot K_2$. Noting that the converse statement, that every member of the set $\{(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')\}$ is a member of the set $\{(\alpha, \beta, \gamma)\}$, is not necessarily true since $b_1 + \cdots + b_m$ may equal zero if $m \geqq 2$, it remains to show that every member of the set $\{(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')\}$ which is likewise a vertex of $K_1 \cdot K_2$ is a member of the set $\{(\alpha, \beta, \gamma)\}$, thus eliminating the possibility that K is actually contained in $K_1 \cdot K_2$. But this follows at once from the fact that each vertex of $K_1 \cdot K_2$ is uniquely determined,[8] i.e., if the point (α, β, γ) is a vertex of K, there is just one term $A'_{\alpha'\beta'\gamma'} x^{\alpha'} y^{\beta'} z^{\gamma'}$ of $F_1(x, y, z)$ and just one term $A''_{\alpha''\beta''\gamma''} x^{\alpha''} y^{\beta''} z^{\gamma''}$ of $F_2(x, y, z)$ such that $(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'') = (\alpha, \beta, \gamma)$. Since $A'_{\alpha'\beta'\gamma'}$ and $A''_{\alpha''\beta''\gamma''} \neq 0$ by hypothesis, $A'_{\alpha'\beta'\gamma'} \cdot A''_{\alpha''\beta''\gamma''} \neq 0$, i.e., $(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')$ is a member of the set $\{(\alpha, \beta, \gamma)\}$. This completes the proof.

From this theorem it follows immediately that for $F(x, y, z)$ to be reducible,[9] it is necessary that its polyhedron decompose.

4. **The general criterion for the case of two variables and one prime.** We next proceed to obtain a criterion of a similar nature for polynomials of two variables with coefficients in the field of rational numbers. Letting $F(x, y, z)$ be such a polynomial, we write it in the form

$$F(x, y) = \sum A_{\alpha\beta\gamma} p^\gamma x^\alpha y^\beta, \quad p \text{ a prime}, \quad A_{\alpha\beta\gamma} \not\equiv 0 \pmod{p}.$$

Before establishing the criterion, let us first examine the nature of the product of two polynomials of the above form. Letting the reducible polynomial

$$F(x, y) = F_1(x, y) \cdot F_2(x, y),$$

[7] m may be equal to 1.

[8] H. Minkowski, loc. cit., p. 181.

[9] Factors which are powers of x and y are excluded, as such a factor corresponds to the excluded case that K has a point factor. Such a factor can, of course, always be determined by inspection.

where $F_1(x, y) = \sum A'_{\alpha'\beta'\gamma'} p^{\gamma'} x^{\alpha'} y^{\beta'}$, $A'_{\alpha'\beta'\gamma'} \not\equiv 0 \pmod p$ and $F_2(x, y) = \sum A''_{\alpha''\beta''\gamma''} p^{\gamma''} x^{\alpha''} y^{\beta''}$, $A''_{\alpha''\beta''\gamma''} \not\equiv 0 \pmod p$, then

$$(1) \qquad F(x, y) = \sum_{\alpha',\beta',\gamma'} \sum_{\alpha'',\beta'',\gamma''} A'_{\alpha'\beta'\gamma'} A''_{\alpha''\beta''\gamma''} p^{\gamma'+\gamma''} x^{\alpha'+\alpha''} y^{\beta'+\beta''}.$$

Now in the expanded product (1) let there be m terms[7] containing $x^\alpha y^\beta$, and let ρ_k be the lowest power of p occurring in these m terms. Setting S equal to the sum of these terms, we have

$$S = (b_1 p^{\rho_i} + \cdots + b_m p^{\rho_j}) x^\alpha y^\beta = A_{\alpha\beta\gamma} p^\gamma x^\alpha y^\beta, \; A_{\alpha\beta\gamma} \not\equiv 0 \pmod p.$$

Noting that if there is just one term containing p^{ρ_k}, $\gamma = \rho_k$, and that if there are two or more terms containing p^{ρ_k}, $\gamma \geqq \rho_k$, we see that unless S drops out, S is always replaced in the contracted product by a single term in which the power of p is at least as large as the lowest power of p occurring in S, i.e., the point (α, β, γ) coincides with or lies above the point (α, β, ρ_k).

With this in mind, we turn to a consideration of the least convex polyhedra K, K_1 and K_2 determined as before by the sets of points $\{(\alpha, \beta, \gamma)\}$, $\{(\alpha', \beta', \gamma')\}$, and $\{(\alpha'', \beta'', \gamma'')\}$, respectively. From the fact that one point (α, β, γ) may replace several different points of the form $(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')$ as indicated above, it is quite clear that the statement $K = K_1 \cdot K_2$ need not be true. To obviate this difficulty, we must replace these finite polyhedra by certain polyhedra infinite in the direction of the positive z-axis. These polyhedra, which we shall term newtonian polyhedra and which we shall denote by N, N_1, and N_2 respectively, are formed by constructing prism-like figures extending upwards indefinitely and having the lower parts of the original polyhedra as bases. A set of points defining any one of these may then be got by adding to the vertices of its base the infinite set of lattice points lying on those of its edges which are parallel to the z-axis. The product of N_1 and N_2 is then defined as before, i.e., if N_1 and N_2 are defined by the sets of points $\{(\alpha_{N_1}, \beta_{N_1}, \gamma_{N_1})\}$ and $\{(\alpha_{N_2}, \beta_{N_2}, \gamma_{N_2})\}$, respectively, their product is defined by the set of points $\{(\alpha_{N_1} + \alpha_{N_2}, \beta_{N_1} + \beta_{N_2}, \gamma_{N_1} + \gamma_{N_2})\}$. By extending Minkowski's work, it then follows that $N_1 \cdot N_2$ as defined above is identical with the newtonian polyhedron formed on the base of $K_1 \cdot K_2$. We are now ready to establish the criterion referred to above, by proving the following

THEOREM 2. *Let the reducible polynomial*

$$F(x, y) = F_1(x, y) \cdot F_2(x, y).$$

Then $N = N_1 \cdot N_2$, where N, N_1 and N_2 are the newtonian polyhedra of F, F_1 and F_2, respectively.

Proof. We may prove this by showing that each vertex of the base of $N_1 \cdot N_2$ is a point of the set $\{(\alpha, \beta, \gamma)\}$ and that the remaining points of the set $\{(\alpha, \beta, \gamma)\}$ lie on or above the base of $N_1 \cdot N_2$.

We start, therefore, by showing that if $A = (\alpha_1, \beta_1, \gamma_1)$ is a vertex of the base of $N_1 \cdot N_2$, it belongs to the set $\{(\alpha, \beta, \gamma)\}$. Since A is a vertex of the base of

$N_1 \cdot N_2$, it is a member of the set $\{(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')\}$, i.e., (1) contains a term $b_i p^{\gamma_1} x^{\alpha_1} y^{\beta_1}$. Further, since A is also a vertex of $K_1 \cdot K_2$, A is determined uniquely, i.e., of the terms of (1) containing $x^{\alpha_1} y^{\beta_1}$, only one contains p^{γ_1}. Moreover, of the powers of p contained in these terms, γ_1 is the smallest, since A belongs to the base of $N_1 \cdot N_2$. Now applying the results obtained above, we see that A must be a member of the set $\{(\alpha, \beta, \gamma)\}$. This completes the first part of the proof.

Turning now to the second part of the proof, we saw above that each point (α, β, γ) coincides with or lies above the point (α, β, ρ_k), where (α, β, ρ_k) is a member of the set $\{(\alpha' + \alpha'', \beta' + \beta'', \gamma' + \gamma'')\}$, i.e., is a point of $K_1 \cdot K_2$ and hence lies on or above the base of $N_1 \cdot N_2$. Therefore, à fortiori, the point (α, β, γ) lies on or above the base of $N_1 \cdot N_2$. This completes the proof.

A direct consequence of this theorem is the necessary condition we sought to establish, namely, that for $F(x, y)$ to be reducible (excluding factors which are powers of x and y as before) it is necessary—but not sufficient—that its newtonian polyhedron decompose.[10]

5. **Special criteria.** We next proceed to combine the general criteria obtained in §§3 and 4 with the necessary conditions for the decomposability of polyhedra obtained in §2 to get certain special classes of algebraic polynomials concerning whose factorability we can formulate definite conclusions.

THEOREM 3. *Let*

$$f(x, y) = a_1\, p^r\, x^{\alpha_1}\, y^{\beta_1} + a_2\, p^s\, x^{\alpha_2}\, y^{\beta_2} + \sum_i a_i\, p^{\gamma_i}\, x^{\alpha_i}\, y^{\beta_i} \quad (i \neq 1, 2)\,,$$

where the line joining the points $A = (\alpha_1, \beta_1, r)$ and $B = (\alpha_2, \beta_2, s)$ has no lattice points, where, further, the points (α_i, β_i) lie on or within a triangle having the line joining the points $a = (\alpha_1, \beta_1)$ and $b = (\alpha_2, \beta_2)$ as a side, and finally, where for all i such that (α_i, β_i) lies on the line ab, i.e., such that $(\alpha_i - \alpha_1)/(\beta_i - \beta_1) = (\alpha_2 - \alpha_1)/(\beta_2 - \beta_1)$,

$$\gamma_i \geqq r + 1 + \left[\frac{s - r}{\alpha_2 - \alpha_1}\,(\alpha_i - \alpha_1)\right].$$

Then, if f has no factor which is a power of x and y, f is irreducible.

Proof. In this case b_{xy} has as one side the line ab. Moreover, the remaining vertex or vertices lie on or within a triangle with ab as a side.[11] Now by hypothesis, a is the projection of a point A of P and b of a point B of P. From the condition on the γ's, it then follows that no points lie directly below AB, i.e., AB

[10] As before, to decompose a newtonian polyhedron means to find two newtonian polyhedra whose product is identical with the original polyhedron. Further, in order that a newtonian polyhedron decompose, it is necessary that its xy-boundary polygon decompose and that the network formed by projecting the lower part of the newtonian polyhedron on the xy-plane decompose.

[11] The reader is asked to construct the figures.

is a side of the lower part of P. Moreover, from the condition that AB has no lattice points, it follows that ab is indecomposable. Hence b_{xy} is indecomposable, and the theorem follows.

In connection with this theorem, it is interesting to note that in its present form it includes as a special case a certain theorem of Glenn[12] for the case of three homogeneous variables. Glenn's theorem[13] applies to a homogeneous polynomial in n variables, and it is easily seen that even our general theorem may be extended to this case. It should be noted, however, that the specialization of our theorem to Glenn's case gives a more precise result, since Glenn's theorem contains an extraneous condition.[14]

THEOREM 4. *Let*

$$f(x, y) = a_1 p^r x^m y^n + a_2 p^s x^{m+i_1} y^{n+j_1} + a_3 p^t x^{m+i_2} y^{n+j_2} + \sum_i a_i p^{\gamma_i} x^{\alpha_i} y^{\beta_i} \qquad (i_1, i_2 \neq 0) ,$$

where each point (α_i, β_i) *lies on or within the parallelogram formed by the lines joining the points* $a = (m, n)$, *and* $d = (m + i_1 + i_2, n + j_1 + j_2)$ *to the points* $b = (m + i_1, n + j_1)$, *and* $c = (m + i_2, n + j_2)$. *Then if* $(r - s, i_1, j_1) = (t - r, i_2, j_2) = 1$, *if, further, for all* i *such that the point* (α_i, β_i) *lies on the line* ab

$$\gamma_i \geqq r + 1 + \left[\frac{s - r}{i_1}(\alpha_i - m)\right],$$

and finally, if for all i *such that the point* (α_i, β_i) *lies on the line* ac

$$\gamma_i \geqq t + 1 + \left[\frac{t - r}{i_2}(\alpha_i - m)\right],$$

[12] O. E. Glenn, *Theorems on reducible quantics*, Annals of Mathematics, (2), vol. 14 (1912–1913), p. 30.

[13] To state his theorem, Glenn first defines normal order as follows: Two sets of p numbers $(k_1, k_2, \cdots, k_p)$, and $(\lambda_1, \lambda_2, \cdots, \lambda_p)$ occur in normal order if the set first to show, when read from right to left, a number greater than the number in the corresponding position in the other set occurs farthest to the right. Then assuming that the terms are arranged so that the subscripts of the coefficients are in normal order, Glenn states the following theorem: A set of necessary conditions that a form f all of whose coefficients within the interval

$$I = \left\{c_{\overbrace{0\cdots0}^{p-\mu}\, m+n\, \overbrace{0\cdots0}^{\mu-1}}, \cdots, {}^*c_{\overbrace{0\quad\cdots\quad0}^{p-\mu-\nu-1}\, m+n\, \overbrace{0\cdots0}^{\mu+\nu}}\right\}$$

are divisible by a prime q be reducible in the absolute field is given by

$$c_{\overbrace{0\cdots0}^{p-i}\, m+n\, \overbrace{0\cdots0}^{i-1}} \equiv 0 \pmod{q^2} \qquad (i = \mu, \mu + 1, \cdots, \mu + \nu).$$

(The starred number is not included in the interval but only shows the upper limit of the interval.)

[14] The extraneous condition referred to is the condition that the coefficients of the terms combining the letters $x_1, x_2, \cdots, x_{p-\mu-\nu-1}$ with the letters $x_{p-\mu-\nu+1}, x_{p-\mu-\nu+2}, \cdots, x_{p-\mu+1}$ be divisible by q, a condition which follows from the fact that these terms are included in the interval I by the definition of normal order. One way of showing that this condition is extraneous is by noting that these terms do not affect the indecomposable side of b_{xy}, and hence f is irreducible even if these terms are *not* divisible by q.

then, factors which are powers of x and y being disregarded, either f is a product of the form

$$f(x, y) = \left(\sum_{\alpha=0}^{i_1} a_\alpha x^\alpha y^{\alpha j_1/i_1}\right)\left(\sum_{\beta=0}^{i_2} a_\beta x^\beta y^{\beta j_2/i_2}\right)$$

(only such values of α and β being taken which make $\alpha j_1/i_1$ and $\beta j_2/i_2$ integral) or f is irreducible. (f may be a product of the indicated form only if f has a term containing $x^{m+i_1+i_2} y^{n+j_1+j_2}$).

Proof. Under these conditions b_{xy} has as two sides the lines ab and ac. Furthermore, the remaining vertex or vertices lie on or within the parallelogram $abdc$. But ab and ac, each being the projection of an indecomposable side of the lower part of P, are likewise both indecomposable. Now let us note that a parallelogram with two adjacent sides indecomposable decomposes uniquely into the product of two line segments, one equal to one of these sides, and the other equal to the other side, and that any other convex polygon having two adjacent sides in common with this parallelogram, and its remaining vertices on the sides or in the interior of the parallelogram, must be indecomposable. It then follows that if b_{xy} has a vertex at $(m + i_1 + i_2, n + j_1 + j_2)$, b_{xy} decomposes uniquely into the product of the two line segments joining $(0, 0)$ to (i_1, j_1) and to (i_2, j_2); otherwise b_{xy} is indecomposable. In the second alternative, of course, f is irreducible. In the first alternative, the factors of P corresponding to the line segments joining $(0, 0)$ to (i_1, j_1) and to (i_2, j_2) are polygons situated entirely in the planes $y = j_1x/i_1$ and $y = j_2x/i_2$, respectively, i.e., f may have factors of the form $\sum_{\alpha=0}^{i_1} a_\alpha x^\alpha y^{\alpha j_1/i_1}$ and $\sum_{\beta=0}^{i_2} a_\beta x^\beta y^{\beta j_2/i_2}$. Then since the decomposition of b_{xy} is unique, factors which are powers of x and y being disregarded, either f is a product of the form

$$f(x, y) = \left(\sum_{\alpha=0}^{i_1} a_\alpha x^\alpha y^{\alpha j_1/i_1}\right)\left(\sum_{\beta=0}^{i_2} a_\beta x^\beta y^{\beta j_2/i_2}\right),$$

or f is irreducible. This completes the theorem.

THEOREM 5. *Let*

$$f(x, y) = a_1 p^r x^{\alpha_1} y^{\beta_1} + a_2 p^s x^{\alpha_2} y^{\beta_2} + \sum_i a_i p^{\gamma_i} x^{\alpha_i} y^{\beta_i} \qquad (i \neq 1, 2),$$

where the line joining $A = (\alpha_1, \beta_1, r)$ to $B = (\alpha_2, \beta_2, s)$ has no lattice points, where the points $a = (\alpha_1, \beta_1)$ and $b = (\alpha_2, \beta_2)$ lie on b_{xy}, where the vertices of b_{xy} on each side of ab lie on or within a triangle with ab as a side, where further the line AB is an edge of P, and where finally for all i such that (α_i, β_i) lies on the line ab

$$\gamma_i \geqq r + 1 + \left[\frac{s - r}{\alpha_2 - \alpha_1}(\alpha_i - \alpha_1)\right].$$

Then, if f has no factor which is a power of x and y, f is irreducible.

Proof. In this case, from the condition on the γ's, it follows that there are no points of P directly below the line AB, i.e., AB is a side or a diagonal of a face

of the lower part of P. But, by hypothesis, AB is an edge, i.e., its projection ab appears in the network formed by projecting the lower part of P. Moreover, since AB is indecomposable, ab is likewise indecomposable. Now let us note that if a network satisfies the conditions (1) an indecomposable side ab of an elementary polygon is a diagonal of its boundary polygon b, and (2) on each side of ab the remaining vertices lie on or within a triangle with ab as side, the network must be indecomposable. It then follows that the network formed by projecting the lower part of P is indecomposable. Hence it follows that, if f has no factor which is a power of x and y, f is irreducible.

We note here that Theorems 3, 4, and 5 are equally valid (with one exception noted below) if the fixed prime p is replaced by the variable z. And, what is more, they are equally valid (with this same exception) if the conditions designed to make certain lines part of the lower part of the polyhedron are replaced by similar conditions designed to make these lines part of the upper part of the polyhedron and p is replaced by z. Thus Theorem 4, for example, is valid (with this same exception) if, in either or both conditions on the γ's, the sign $\geqq$ is replaced by the sign $<$, when p is replaced by z. The exception mentioned is that $f(x, y, z)$ may have a factor of the form $a_1 + a_2 z + \cdots + a_r z^r$. This possibility arises from the fact that the indecomposability of b_{xy} does not preclude the possibility of a point factor of b_{xy}. But such a point factor of b_{xy} may be the projection of a line factor of the polyhedron, parallel to the z-axis, and may thus give rise to a factor of $f(x, y, z)$ of the form indicated. However, the existence or non-existence of such a factor can always be determined by elementary methods without resort to polyhedra.

6. **The case of one variable and two primes.** We conclude this paper by a few remarks on the treatment by Fujiwara[15] of a problem of a similar nature, namely, the application of the polyhedra method to polynomials in one variable written in the form $f(x) = \sum A_{\alpha\beta\gamma}\, p^\alpha\, q^\beta\, x^\gamma$, p and q primes and $A_{\alpha\beta\gamma} \not\equiv 0 \pmod{p, q}$. After defining $K'(f)$ to be that part of the surface of the polyhedron of $f(x)$ for which the direction of the inner normals lies within the domain $(x \geqq 0, y \geqq 0)$ and assuming that $h(x) = f(x)\cdot g(x)$, Fujiwara states the following result:[16] If $K'(f)$ and $K'(g)$ have no parallel faces, then $K'(h)$ contains every face of $K'(f)$ and $K'(g)$, and only these faces, unchanged as to shape and direction and only changed as to position; and furthermore, that if $K'(f)$ and $K'(g)$ have a pair of parallel faces $\pi(f)$ and $\pi(g)$, then $K'(h)$ has a face $\pi(h)$ parallel to these faces and the boundary of $\pi(h)$ is the product polygon of the boundaries of $\pi(f)$ and $\pi(g)$. By the application of these results to polynomials $f(x)$ so constructed that $K'(f)$ shall consist only of a single triangle, Fujiwara

[15] M. Fujiwara, *Über Kriterien für Irreduzibilität ganzzahliger algebraischer Gleichungen*, Tôhoku Mathematical Journal, vol. 17 (1920), pp. 10–17.

[16] Loc. cit., p. 14.

then derives the result[17] that

$$f(x) = a_0x^5 + a_1\,pq^3\,x^4 + a_2\,pq^7\,x^3 + a_3\,p^4q^5\,x^2 + a_4\,p^6q^5\,x + a_5\,p^8q^5\,,$$

where all the a's $\not\equiv 0 \pmod{p,\,q}$, is irreducible.

However, in view of the fact that the congruence properties of p and q are entirely independent, it appeared to us à priori unlikely that results of so general a nature could be true. Indeed, by taking $f(x) = a_0x^7 + a_1p^3qx^5 + a_2p^6x^2 + a_3p^4q^3$ and $g(x) = b_0x^5 + b_1p^3qx^3 + b_2q^3x^2 + b_3p^5q^2$, where the a's and b's are all $\not\equiv 0 \pmod{p,\,q}$, and assuming that $a_1b_0 + a_0b_1 \not\equiv 0 \pmod{p,\,q}$, we find that $K'(h)$ does not contain the faces of $K'(f)$ or of $K'(g)$, in contradiction to Fujiwara's general result above. In fact, this example illustrates the more or less complete breakdown of the polyhedra method in this case, for not one of the faces of the polyhedron of $f(x)$ or of $g(x)$ appears in the polyhedron of $h(x)$. Furthermore, if we now take

$$f(x) = 40{,}755{,}504x^5 - 26\cdot 5\cdot 11^3x^4 - 696\cdot 5\cdot 11^7x^3 - 5^4\cdot 11^5x^2 - 5^6\cdot 11^5x - 3{,}261\cdot 5^8\cdot 11^5\,,$$

we find that $f(x)$ is reducible into the product of $f'(x)$ and $f''(x)$, where

$$f'(x) = 40{,}755{,}504x^3 - 1{,}630{,}346\cdot 5^3\cdot 11x^2 + 81{,}514\cdot 5^5\cdot 11^2x - 3{,}261\cdot 5^7\cdot 11^3$$

and

$$f''(x) = x^2 + 5\cdot 11x + 5\cdot 11^2\,,$$

in contradiction to Fujiwara's special criterion above.

In conclusion, we wish to state that whether or not the above results can be supplemented by results of a positive nature for this case remains to be determined. However, the above considerations seem to indicate that there is little, if any, possibility of extending the polyhedra method to this case.

YALE UNIVERSITY.

[17] Loc. cit., p. 17.

FOURIER SERIES CONVERGENCE CRITERIA, AS APPLIED TO CONTINUOUS FUNCTIONS

By J. A. Clarkson and W. C. Randels

Whether or not there exists a continuous function whose Fourier series diverges everywhere, or almost everywhere, or on a set of points of positive measure, remains an unsolved problem. If a local condition is known which is sufficient to insure convergence of the Fourier series at a point, one is naturally led to raise the same question about the local condition itself: do there exist continuous functions which violate it at every point? For the criterion of Jordan, for example, the answer is clearly yes; for the more recent and more delicate criteria the question presents greater difficulty. Mazurkiewicz[1] and Kaczmarz[2] have shown that the answer is also affirmative in the case of the Dini criterion. It is the purpose of this note to answer this question for several more general convergence criteria.

Given any continuous function $f(x)$, which is periodic with period 2π, we define

$$\varphi(f; x; t) = \varphi(t) = f(x + t) + f(x - t) - 2f(x)$$

and

$$\Delta_\delta^k \varphi(t) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi(t + j\delta).$$

We first consider the condition

$$(L_k) \qquad \lim_{\delta \to +0} \int_\delta^\pi \frac{1}{t} \,|\, \Delta_\delta^k \varphi(t) \,|\, dt = 0$$

for a fixed integer k; any condition L_k insures convergence of the Fourier series, and the conditions are increasingly general; that is, L_k implies L_{k+1}. L_1 is the familiar Lebesgue criterion.

Let C be the space of continuous functions, periodic (2π), with the customary norm. We first prove

Theorem 1. *For any positive integer k, the subset $A \subset C$ of functions such that for each x we have*

$$\overline{\lim_{\delta \to +0}} \int_\delta^\pi \frac{1}{t} \,|\, \Delta_\delta^k \varphi(t) \,|\, dt = +\infty$$

is of the second category in C, and its complement is of the first category.

Received January 8, 1935 by the Editors of the Annals of Mathematics, accepted by them and later transferred to this journal. The second named author is Sterling Research Fellow at Yale University.

[1] Studia Mathematica, vol. 3 (1931), p. 114.

[2] Ibid., p. 189.

We shall show that Theorem 1 is a consequence of the following theorem of Banach:[3]

THEOREM A. [Banach]. *Let the operation* $U(f, x, \delta)$, *which makes correspond to every element* f *of a Banach space* B, *every number* x $(-\pi \leqq x < \pi)$ *and every number* δ $(0 < \delta \leqq 1)$ *a real non-negative number* U, *satisfy the following conditions:*

(i) *For* δ *fixed,* $U(f, x, \delta)$ *is continuous in* f *and* x.

(ii)
$$U(f, x, \delta) = U(-f, x, \delta),$$
$$U(f + g, x, \delta) \leqq U(f, x, \delta) + U(g, x, \delta).$$

(iii) *There exists an everywhere dense set* $H \subset B$ *such that for every* $\omega \in H$, $U(\omega, x, \delta)$ *is a bounded function of* x *and* δ.

(iv) *Given* r, $M > 0$, *there exists an element* $g \in B$, $\| g \| < r$, *with*
$$\sup_{\delta} U(g, x, \delta) > M \quad (-\pi \leqq x < \pi).$$

The set D, *of elements* f *in* B *such that for all* x
$$\overline{\lim_{\delta \to +0}}\, U(f, x, \delta) = +\infty,$$
is of the second category in B, *and a complement of a set of the first category.*

Let C be the Banach space in question; clearly the operation
$$U(f, x, \delta) = \int_{\delta}^{\pi} \frac{1}{t} \,|\, \Delta_{\delta}^{k} \varphi(t) \,|\, dt$$
satisfies conditions (i) and (ii) of Theorem A, and condition (iii) is satisfied by taking for H the set of trigonometric polynomials. Since $U(cf, x, \delta) = |\, c \,|$ $U(f, x, \delta)$, condition (iv) will be satisfied if there exists in C a sequence $\{g_n(x)\}$ with $\| g_n \| \leqq 1$, and $\lim_{n \to \infty} \left[\inf_{x} \sup_{\delta} U(g_n, x, \delta) \right] = +\infty$. We proceed to show that $g_n(x) = |\sin nx|$ is such a sequence.

By virtue of the periodicity (π/n) in x of $U(g_n, x, \delta)$, we need only establish that
$$\lim_{n \to \infty} \left[\inf_{0 \leqq x \leqq \pi/n} \sup_{\delta} U(g_n, x, \delta) \right] = +\infty.$$

Let x_0 be that solution of $2 \sin x - \sqrt{2} \cos x = 0$ which lies in the interval $(0, \pi/4)$. Let I_ν^n $(\nu = 1, 2, \cdots, 8;\ n = 1, 2, 3, \cdots)$ be the intervals $(0, x_0/n)$, $(x_0/n, \pi/4n)$, $[\pi/4n, (\pi - 2x_0)/2n]$, $[(\pi - 2x_0)/2n, \pi/2n]$, etc. We shall show that corresponding to each ν there are three constants a_ν, δ_ν, c_ν $(0 \leqq a_\nu \leqq \pi/4$, $0 < \delta_\nu \leqq \pi/2$, $c_\nu > 0)$ such that[4] for $x \in I_\nu^n$,
$$|\, \Delta_{\delta_\nu/n}^{k} \varphi_n(a_\nu/n) \,| > c_\nu \qquad (n = 1, 2, 3, \cdots).$$

If we assume this to be true, since $\frac{d}{dt} \{\Delta_{\delta_\nu/n}^{k} \varphi_n(t)\} = O(n)$ uniformly in x and t, it will follow that there will exist an interval (a_ν', a_ν'') containing a_ν, of length

[3] Banach, *Über die Bairesche Kategorie gewisser Functionenmengen*, Studia Mathematica, vol. 3 (1931), pp. 174–179. The changes in statement are non-essential.

[4] In the following we employ the notation $\varphi(g_n, x, t) = \varphi_n(t)$.

h_ν, such that for $a'_\nu/n \leqq t \leqq a''_\nu/n$ and $x \in I^n_\nu$ we have $| \Delta^k_{\delta_\nu/n} \varphi_n(t) | > c_\nu/2$ $(n = 1, 2, 3, \cdots)$. There is such an interval within each interval

$$[m\pi/n, (m+1)\pi/n] \qquad (m = 1, 2, \cdots, n-1),$$

whence we have, for all $x \in I^n_\nu$,

$$U(g_n, x, \delta_\nu/n) = \int_{\delta_\nu/n}^{\pi} \frac{1}{t} | \Delta^k_{\delta_\nu/n} \varphi_n(t) | \, dt \geqq \sum_{m=1}^{n-1} \int_{m\pi/n}^{(m+1)\pi/n} \frac{1}{t} | \Delta^k_{\delta_\nu/n} \varphi_n(t) | \, dt$$

$$\geqq \sum_{m=1}^{n-1} \frac{c_\nu}{2} \cdot \frac{n}{(m+1)\pi} \cdot \frac{h_\nu}{n} = \frac{c_\nu h_\nu}{2\pi} \sum_{m=1}^{n-1} \frac{1}{m+1}.$$

Thus, uniformly for $x \in I^n_\nu$, $U(g_n, x, \delta_\nu/n) \to \infty$ as $n \to \infty$, and the desired property of the sequence will be established. It merely remains, then, to show that the constants a_ν, δ_ν, c_ν can be chosen as stated.

For $x \in I^n_1$, we choose $a_1 = 0$, $\delta_1 = \pi/2$. Then

$$\Delta^k_{\pi/2n} \varphi_n(0) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi_n(j\pi/2n)$$

$$= \sum_{j=0}^{\left[\frac{k}{2}\right]} \binom{k}{2j} \varphi_n(0) - \sum_{j=0}^{\left[\frac{k-1}{2}\right]} \binom{k}{2j+1} \varphi_n(\pi/2n)$$

$$= \frac{1}{2} [\varphi_n(0) - \varphi_n(\pi/2n)] \sum_{j=0}^{k} \binom{k}{j} = (\cos nx - \sin nx) \sum_{j=0}^{k} \binom{k}{j}.$$

This clearly exceeds some $c_1 > 0$ for $x \in I^n_1$.

For $x \in I^n_2$ we take $a_2 = \delta_2 = \pi/4$. Then we have

$$\Delta^k_{\pi/4n} \varphi_n(\pi/4n) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \varphi_n\left(\frac{\pi}{4n} + \frac{j\pi}{4n}\right)$$

$$= \sum_{j=0}^{\left[\frac{k}{4}\right]} \binom{k}{4j} \varphi_n(\pi/4n) - \sum_{j=0}^{\left[\frac{k-1}{4}\right]} \binom{k}{4j+1} \varphi_n(\pi/2n)$$

$$+ \sum_{j=0}^{\left[\frac{k-2}{4}\right]} \binom{k}{4j+2} \varphi_n(3\pi/4n) - \sum_{j=0}^{\left[\frac{k-3}{4}\right]} \binom{k}{4j+3} \varphi_n(0)$$

$$= (\sqrt{2} \cos nx - 2 \sin nx) \sum_{j=0}^{\left[\frac{k}{2}\right]} \binom{k}{2j} - 2(\cos nx - \sin nx) \sum_{j=0}^{\left[\frac{k-1}{4}\right]} \binom{k}{4j+1}$$

$$< (\sqrt{2} - 2) \cos nx \sum_{j=0}^{\left[\frac{k-1}{4}\right]} \binom{k}{4j+1}.$$

This again is bounded away from zero uniformly for $x \in I_2^n$.

The remaining six cases are handled in similar fashion. We omit the computation, but supply the following table showing, for each value of ν, the values of a_ν and δ_ν and the value of $|\Delta^k_{\delta_\nu/n}\, \varphi_n(a_\nu/n)|$ or a function which is less than the latter in absolute value as in the second case above.

Interval	a_ν	δ_ν	$\lvert \Delta^k_{\delta_\nu/n}\, \varphi_n(a_\nu/n) \rvert$
$0,\ x_0/n$	0	$\pi/2$	$= (\cos nx - \sin nx) \sum_{j=0}^{k} \binom{k}{j}$
$x_0/n,\ \pi/4n$	$\pi/4$	$\pi/4$	$\geqq (2 - \sqrt{2}) \cos nx \sum_{j=0}^{\left[\frac{k-1}{4}\right]} \binom{k}{4j+1}$
$\pi/4n,\ (\pi - 2x_0)/2n$	$\pi/4$	$\pi/4$	$\geqq (2 - \sqrt{2}) \sin nx \sum_{j=0}^{\left[\frac{k-1}{4}\right]} \binom{k}{4j+1}$
$(\pi - 2x_0)/2n,\ \pi/2n$	0	$\pi/2$	$= (\sin nx - \cos nx) \sum_{j=0}^{k} \binom{k}{j}$
$\pi/2n,\ (\pi + 2x_0)/2n$	0	$\pi/2$	$= (\sin nx + \cos nx) \sum_{j=0}^{k} \binom{k}{j}$
$(\pi + 2x_0)/2n,\ 3\pi/4n$	$\pi/4$	$\pi/4$	$\geqq (2 - \sqrt{2}) \sin nx \sum_{j=0}^{\left[\frac{k-1}{4}\right]} \binom{k}{4j+1}$
$3\pi/4n,\ (\pi - x_0)/n$	$\pi/4$	$\pi/4$	$\geqq (\sqrt{2} - 2) \cos nx \sum_{j=0}^{\left[\frac{k-1}{4}\right]} \binom{k}{4j+1}$
$(\pi - x_0)/n,\ \pi/n$	0	$\pi/2$	$= -(\sin nx + \cos nx) \sum_{j=0}^{k} \binom{k}{j}$

The sequence $\{g_n(x)\}$ has, then, the required property; Theorem A may be applied, and our result follows at once.

Gergen[5] has given a complete analysis of the various known criteria for convergence of a Fourier series, and suggested certain generalizations. In particular he shows that a sufficient condition is given by

$$(G_k) \qquad \lim_{\xi \to +\infty} \overline{\lim_{\delta \to 0}} \int_{\xi\delta}^{\pi} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt = 0 .$$

As in the case of the L_k, G_k implies G_{k+1}, and G_1 is implied by all of the previously known criteria mentioned by Gergen. Our theorem allows us to infer at once

THEOREM 2. *For any positive integer k, the set $D \subset C$ of functions f such that for each x we have*

$$\lim_{\xi \to +\infty} \overline{\lim_{\delta \to 0}} \int_{\xi\delta}^{\pi} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt = +\infty$$

is of the second category in C, and its complement is a set of the first category.

Proof. By virtue of Theorem 1, it will be sufficient to show that at a fixed point x the condition

$$(1) \qquad \overline{\lim_{\delta \to +0}} \int_{\delta}^{\pi} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt = +\infty$$

implies

$$\lim_{\xi \to +\infty} \overline{\lim_{\delta \to +0}} \int_{\xi\delta}^{\pi} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt = +\infty .$$

Assume (1) to be true; then for any fixed $\xi > 1$ we have

$$\int_{\delta}^{\pi} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt - \int_{\xi\delta}^{\pi} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt = \int_{\delta}^{\xi\delta} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt ,$$

which, as $\mid \Delta_\delta^k \varphi(t) \mid$ is less than some bound B, is

$$< B \int_{\delta}^{\xi\delta} \frac{dt}{t} = B \log \xi .$$

Thus for any fixed ξ, $\overline{\lim}_{\delta \to +0} \int_{\xi\delta}^{\pi} \frac{1}{t} \mid \Delta_\delta^k \varphi(t) \mid dt = +\infty$, and Theorem 2 follows.

It may also be noted that the total additivity of the set property of being of the first category allows us to combine the above results and state finally

THEOREM 3. *The set $E \subset C$ of functions which at no point x satisfy any of the conditions G_k or L_k ($k = 1, 2, \cdots$) is of the second category in C, and its complement is of the first category.*

INSTITUTE FOR ADVANCED STUDY AND YALE UNIVERSITY.

[5] Gergen, *Convergence and summability criteria for Fourier series*, Quarterly Journal of Mathematics, (Oxford Series), vol. 1 (1930), p. 252.

ON LOCAL BETTI NUMBERS

By H. E. Vaughan, Jr.

1. Introduction. Several types of local Betti numbers have been introduced recently by Alexandroff[1] and by Čech.[2] The local invariants introduced in this paper were discovered during an attempt to define edge and kernel points of a compact metric space. Incidentally, they give a direct generalization of the notion of the order, at a point, of a 1-dimensional set.[3]

Section 2 consists of a list of theorems, a knowledge of which is necessary in the later sections. In §3 the numbers $\beta^i(a, M)$, $i \geqq 0$, are defined for each point a of a compact metric space M, and this definition is illustrated in §4 by examples. In §5 are given several definitions of edge and kernel points which lead to simple necessary conditions that a compact metric space be imbeddable in the compact euclidean space of the same dimension. §6 is devoted to the determination of the Borel class of the set of all points of M for which the numbers $\beta^i(a, M)$ satisfy certain inequalities.

In §7 the numbers $\beta^i(a, M)$ are related to the local connectedness of the set M, and also to that of its complement when M is considered as a subset of a euclidean space. In order to extend these theorems, certain auxiliary theorems on the addition of irreducible membranes are required, and these are given in §8. Their immediate consequences are then developed in §9.

There exist in the literature numerous characterizations of the plane, the closed 2-cell and 2-manifold. The majority of these are purely set-theoretic, excepting certain definitions of Whitney and van Kampen[4] which make use of mixed methods. We give below, in §10, a characterization of the 2-manifold in terms of the numbers $\beta^i(a, M)$. In §11 it is shown that a similar characterization can be given for the closed 2-cell and, in fact, for any 2-dimensional set obtained from a 2-manifold by the omission of a finite number of open 2-cells. In §12 necessary and sufficient conditions are given that every point of a locally compact metric space have a neighborhood homeomorphic with a 2-cell, and these are applied to give characterizations of the open 2-cell (or euclidean plane) and of the class of cylinder-trees.[5] The characterizations mentioned in this paragraph are of a purely combinatorial nature.

Received June 17, 1935; presented to the American Mathematical Society, April 19, 1935.

[1] *On local properties of closed sets*, Annals of Mathematics, vol. 36 (1935), pp. 1–35.

[2] *Sur les nombres de Betti locaux*, Annals of Mathematics, vol. 35 (1934), pp. 678–701.

[3] Menger, *Kurventheorie*, p. 96.

[4] van Kampen, *On some characterizations of 2-dimensional manifolds*, this journal, vol. 1 (1935), p. 87.

[5] Zippin, *On continuous curves and the Jordan curve theorem*, American Journal of Mathematics, vol. 52 (1930), pp. 331–350.

In §13 some properties of Alexandroff's local Betti numbers are proved and an inequality is shown to exist between these and the local invariants of the present paper.

Section 14 contains a list of unsolved problems.

I wish to take this opportunity to express my indebtedness to Professor R. L. Wilder who has supervised this investigation and who has aided me constantly by giving many valuable suggestions.

2. Theorems used in the succeeding sections. In this paper R^n will always denote the compact n-dimensional euclidean space.

Definition. If A is any metric set and S is any system of open subsets of A, the S-regular part of A is the set of all points of A each of which is contained in arbitrarily small sets of the system S. The S-irregular part of A is the complement with respect to A of the S-regular part.

Theorem A. For every metric set A and every system S of open subsets of A, the S-regular part of A is a G_δ in A, the S-irregular part of A an F_σ in A.[6]

Theorem B. If M is a compact metric space and $p^i(M)$ is finite, there exists an $\eta > 0$ such that every complete i-cycle of diameter $< \eta$ bounds on M. If M is a compact metric space which is locally j-connected, $0 \leqq j \leqq i$, then $p^i(M)$ is finite.[7]

Theorem C. Let C be the sum and C^{i-1} the intersection of two closed sets of points, A and B, in R^n. Then every k-cycle L^k, $k < n - 1$, of $R^n - C$ which bounds a chain L_A^{k+1} of $R^n - A$ and a chain L_B^{k+1} of $R^n - B$ must also bound in $R^n - C$ provided the chains L_A^{k+1} and L_B^{k+1} may be so chosen that $L_A^{k+1} + L_B^{k+1}$ bounds in $R^n - C^{i-1}$. This is true even for $k = n - 1$, unless C^{i-1} is vacuous.[8]

Theorem D. Let F', F'' be two closed subsets of R^m such that $F'F''$ carries a complete r-cycle which fails to bound on $F'F''$ but which bounds on F' and on F''. There exists an $(m - r - 2)$-cycle in $R^m - (F' + F'')$ which bounds in $R^m - F'$ and in $R^m - F''$, but not in $R^m - (F' + F'')$.[9]

Theorem E. Let F', F'' be two closed subsets of R^m, and γ^{m-r-2} a cycle in $R^m - (F' + F'')$ which bounds in $R^m - F'$ and in $R^m - F''$ but not in $R^m - (F' + F'')$. Then $F'F''$ carries a complete r-cycle which bounds on F' and on F'' but not on $F'F''$.[9]

Theorem F. Let M be a compact metric space which is the irreducible carrier of an essential complete m-cycle and K a closed subset of M such that $p^{m-1}(K) = k$. Then $M - K$ has at most $k + 1$ components.[10]

Theorem G. A locally 0-connected compact metric space is homeomorphic

[6] Menger, *Kurventheorie*, p. 103.

[7] Wilder, *On locally connected spaces*, this journal, vol. 1 (1935), pp. 543–555.

[8] Alexander, *A proof and extension of the Jordan-Brouwer separation theorem*, Transactions of the American Mathematical Society, vol. 23 (1922), p. 342.

[9] Alexandroff, *Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension*, Annals of Mathematics, vol. 30 (1928), p. 178.

[10] Wilder, *Domains and their boundaries in E_n*, Mathematische Annalen, vol. 109 (1933), p. 281.

with a 2-dimensional manifold if it contains irreducibly a 2-cycle and is separated by each simple closed curve of diameter less than $\delta > 0$.[4]

THEOREM H. A necessary and sufficient condition that a locally 0-connected, locally compact metric continuum be a cylinder-tree is that it be cut by every simple closed curve but by no arc.[11]

THEOREM I. A necessary and sufficient condition that a locally 0-connected, locally compact metric continuum S be a cylinder-tree is that it be cyclically connected and, if K is any simple closed curve of S, every point of K is a limit point of $S - K$ and $S - K$ is the sum of precisely two components.[12]

THEOREM J. A necessary and sufficient condition that a locally 0-connected, locally compact cyclically connected metric continuum be homeomorphic with a subset of a spherical surface is that it do not contain a primitive skew curve.[13]

3. **Definition of $\beta^i(a, M)$.** Let M be a compact metric space, a a point of M, and $k = \dim_a M$. There exist arbitrarily small neighborhoods of a whose boundaries are $(k - 1)$-dimensional compact metric spaces. For every non-negative integer i and real number $\epsilon > 0$, let $\beta^i_\epsilon(a, M)$ be the smallest integer b such that there exists a neighborhood G of a such that $\delta(G) < \epsilon$, $\dim(\bar{G} - G) = k - 1$, and[14] $p^i(\bar{G} - G) = b$. Then $\beta^i_\epsilon(a, M)$ is defined for every $\epsilon > 0$ and, as ϵ approaches zero, is a monotone, non-decreasing function. Consequently it approaches a limit, which is a non-negative integer or ∞. In case the limit is finite it is denoted by $\beta^i(a, M)$. If the limit is infinite, two cases arise: (1) $\beta^i_\epsilon(a, M)$ is finite for all $\epsilon > 0$, in which case $\beta^i(a, M) = \omega$; (2) for sufficiently small values of ϵ, $\beta^i_\epsilon(a, M) = \infty$, in which case $\beta^i(a, M) = \aleph_0$. In the definition of $\beta^0(a, M)$ a 0-cycle is defined as an even number of points. The coefficient domain is, of course, arbitrary, but in the present paper it will always be assumed finite, i.e., mod $m \geqq 2$, for reasons of convergence.

Remarks. I. From the fundamental properties of complete i-cycles it follows that $\beta^i(a, M) = 0$ for $i > k - 1$.

II. It is immediately evident from the definition that $\beta^i(a, M)$ is a local topological invariant of M and, in particular, is independent of any space in which M may be considered to be imbedded.

III. Although the hypothesis that the $\beta^i(a, M)$ have definite values makes it possible to choose, for each value of i, an arbitrarily small neighborhood G satisfying the conditions $\dim(\bar{G} - G) = k - 1$, $p^i(\bar{G} - G) = \beta^i(a, M)$, if this number is finite, or $p^i(\bar{G} - G)$ finite if $\beta^i(a, M) = \omega$, it is not in general possible, as can be shown by examples, to choose a single neighborhood G which satisfies these

[11] Zippin, loc. cit., p. 341.

[12] Zippin, loc. cit., p. 348.

[13] Claytor, *Topological immersion of peanian continua in a spherical surface*, Annals of Mathematics, vol. 35 (1934), p. 832, and Zippin, *On semi-compact spaces*, American Journal of Mathematics, vol. 57 (1935), p. 339.

[14] Vietoris, *Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen*, Mathematische Annalen, vol. 97 (1926), pp. 454–472. When we speak of complete cycles we refer to the Vietoris "Fundamentalfolgen".

conditions for all values of i. Of course, by a proper modification of the definition, this could be done, but as yet it has not appeared desirable to add this extra complication.

IV. If the set M is considered as lying in a second space, R^n, for instance, the definition of $\beta^i(a, M)$ can be given in terms of neighborhoods of a in this space. It has not, however, been shown, and indeed seems unlikely, that similar topological invariants can be defined in terms of any particular class of neighborhoods in the imbedding space. This is due to the impossibility, in general, of extending a topological transformation. It can be easily shown that definition in terms of spherical neighborhoods would not give topological invariants. In this connection it is to be noted that while the local invariants introduced by Alexandroff are defined in terms of spherical neighborhoods, a double limit process is required.

4. **Examples.** 1. For any point a of a compact metric space such that $\dim_a M = 1$, we have $\beta^0(a, M) = \mathrm{ord}_a M - 1$, if the order is finite, $\beta^0(a, M) = \mathrm{ord}_a M$, if the order is ω or $\aleph_0$, and $\beta^0(a, M) = \aleph_0$, if the order is $\aleph_0$ or c. $\beta^i(a, M) = 0$, if $i > 0$.

2. For any point a of an n-dimensional (combinatorial) manifold, or interior point of an n-cell, $\beta^i(a, M) = 0$, $(0 \leqq i < n - 1)$, and $\beta^{n-1}(a, M) = 1$. For any boundary point a of a (closed) n-cell $\beta^i(a, M) = 0$, $i \geqq 0$.

3. Let M be a set consisting of n 2-cells with a common edge. If a is a point of this edge, then $\beta^0(a, M) = 0$, $\beta^1(a, M) = 0$ or $n - 1$ according as a is or is not an end point, and $\beta^i(a, M) = 0$, if $i > 1$.

4. Let M be a set consisting of n 2-cells having only an interior point a in common. Then $\beta^0(a, M) = n - 1$, $\beta^1(a, M) = n$ and $\beta^i(a, M) = 0$, if $i > 1$.

5. **Definition of edge point and kernel point.** The usual definitions of boundary point and interior point of a set M imbedded in R^n are as follows. The point a is a boundary point of M if every sphere with center at a contains a point of M and a point not of M. The point a is an interior point of M if there exists a sphere with center at a which is entirely contained in M. These definitions are taken to define edge points and kernel points respectively in the case of an n-dimensional closed subset of R^n, and the purpose of this section is to give them an invariant formulation in terms of the local Betti numbers.

First, if M is an n-dimensional closed subset of R^n, $n \neq 0$, and $\beta^{n-1}(a, M) = 0$, then a is a boundary point of M. For, since M is not 0-dimensional, a is a limit point of M, while inside any sphere with center at a there is a neighborhood (in R^n) whose boundary intersects M in a set whose $(n - 1)$-th Betti number is zero and hence is a proper subset of this boundary. Consequently, interior to the sphere there is a point of this boundary not belonging to M. This proves the statement.

Conversely, if a is a boundary point of M, $\beta^{n-1}(a, M) = 0$. For there exists an arbitrarily small sphere having a as center whose boundary is not contained in M.

These remarks lead to the following definitions, which will be further justified later.

DEFINITION. The point a is called a *k-edge point* of the compact metric space M if $\dim_a M = k$ and $\beta^{k-1}(a, M) = 0$.

DEFINITION. The point a is called a *k-kernel point* of the compact metric space M if $\dim_a M = k$ and $\beta^{k-1}(a, M) > 0$.

DEFINITION. The point a is called an *ordinary k-kernel point* of the compact metric space M if $\dim_a M = k$ and $\beta^{k-1}(a, M) = 1$.

DEFINITION. The point a is called a *regular k-kernel point* of the compact metric space M if $\dim_a M = k$ and $\beta^i(a, M) = 0$, $0 \leqq i < k - 1$, $\beta^{k-1}(a, M) = 1$.

The preceding discussion shows that a necessary condition that it be possible to imbed a compact metric space M in R^n is that $\beta^{n-1}(a, M) \leqq 1$ for every point a of M, while for every point such that $\beta^{n-1}(a, M) = 1$ it is necessary that $\beta^i(a, M) = 0$, $0 \leqq i < n - 1$. This may be restated in the following

THEOREM 1. *A necessary condition that a compact n-dimensional metric space M be imbeddable in R^n is that every point a satisfying* $\dim_a M = n$ *be either an n-edge point or a regular n-kernel point.*

The above condition is naturally not sufficient. In fact, it is not even sufficient for "local imbeddability", as may be shown by the example of a sphere with infinitely many "handles". In this case a point a exists such that no neighborhood of a can be imbedded in R^2. Moreover, it is possible to construct a 2-dimensional set, every point of which is a regular 2-kernel point but which contains no open subset which can be imbedded in R^2.

Several other definitions of kernel points and edge points have been given, including two by Alexandroff.[15] By a result of his[15] it follows that every n-dimensional set contains an n-kernel point as defined above.

6. **Closure properties of certain sets.** It is possible to apply Theorem A to the solution of this problem. To do so, let S_1 be the class of all open subsets G of M such that $\dim(\bar{G} - G) \leqq k - 1$, let S_2 be the class of all open subsets G of M such that $\dim(\bar{G} - G) \leqq k - 1$, $p^i(\bar{G} - G) \leqq p$, let S_3 be the class of all open subsets G of M such that $\dim(\bar{G} - G) \leqq k - 1$, $p^i(\bar{G} - G)$ finite. Using these for S in Theorem A and letting $n = \dim M$, $k \leqq n$, the following results are obtained.[16]

1. The set of points a of M for which

$$\begin{array}{lll} \dim_a M & \leqq k & \text{is a } G_\delta \\ & > k & F_\sigma \\ & < k & G_\delta \\ & \geqq k & F_\sigma \\ & = k & G_{\delta\rho},\ G_{\delta\sigma},\ F_{\sigma\delta} \\ & = n & F_\sigma. \end{array}$$

(These relations, due to Menger, are well known.)

[15] See footnote 1, p. 27, and *Dimensionstheorie*, Mathematische Annalen, vol. 106 (1932), pp. 161–238.

[16] If A represents a class of sets, the symbol A_ρ represents the class consisting of those sets which may be obtained as the difference of two sets of the class A. See Menger, *Kurventheorie*, p. 105.

2. Those points a of the S_2-regular part of M for which $\dim_a M = k$ have $\beta^i(a, M) \leqq p(< p + 1)$, and the S_2-regular part of M contains only points a such that $\dim_a M \leqq k$. Those points a of the S_2-irregular part of M for which $\dim_a M = k$ have $\beta^i(a, M) > p(\geqq p + 1)$.

3. Those points a of the S_3-regular part of M for which $\dim_a M = k$ have $\beta^i(a, M)$ finite or ω, and the S_3-regular part of M contains only points a such that $\dim_a M \leqq k$. Those points a of the S_3-irregular part of M for which $\dim_a M = k$ have $\beta^i(a, M) = \aleph_0$.

From these remarks several results may be deduced, of which the following are the more important.

The set of points a of M such that $\dim_a M = k$ and

$$\begin{array}{lll} \beta^i(a, M) & \leqq \aleph_0 \text{ form a} & G_{\delta\rho} \\ & = \aleph_0 & G_{\delta\rho\rho} \\ & \geqq \omega & G_{\delta\rho} - G_{\delta\rho\sigma} \\ & = \omega & G_{\delta\rho} - G_{\delta\rho\sigma} \\ & \leqq \omega & G_{\delta\rho} \\ & < \omega & G_{\delta\rho\sigma} \\ & > p & G_{\delta\rho\rho} \\ & \geqq p & G_{\delta\rho\rho} \\ & = p & G_{\delta\rho\rho} \\ & \leqq p & G_{\delta\rho} \\ & < p & G_{\delta\rho}. \end{array}$$

THEOREM 2. *The set of k-edge points of a compact metric space is a* $G_{\delta\rho}$.

THEOREM 3. *The set of k-kernel points of a compact metric space is a* $G_{\delta\rho\rho}$.

THEOREM 4. *The set of ordinary k-kernel points of a compact metric space is a* $G_{\delta\rho\rho}$.

7. Local connectedness. In the examples of §4 we have seen that the local Betti numbers give a measure of the ramification of the compact metric space M. In the present section we give some theorems relating the numbers $\beta^i(a, M)$ to the local i-connectedness properties of M, and, in case M is imbedded in R^n, to the uniform local i-connectedness of $R^n - M$.

DEFINITION. If a is a point of the compact metric space M such that to every $\epsilon > 0$ there corresponds a $\delta > 0$ such that every complete i-cycle carried by $S(a, \delta)$ bounds on $S(a, \epsilon)$, then M is said to be *locally i-connected at the point a*. If M is locally i-connected at each of its points it is said to be *locally i-connected*.

The following theorem shows the relation between local i-connectedness and the local Betti numbers.

THEOREM 5. *Let M be a compact metric space and a a point of M such that* $\beta^i(a, M)$ *is finite or* ω. *Suppose further that one of the following three conditions is satisfied.*

1. $p^i(M)$ *is finite.*

2. *There exists a real number* $\eta > 0$ *such that every complete i-cycle carried by* $S(a, \eta)$ *bounds on* M.

3. $p^i(a, M) = 0$.[17]
Then M is locally i-connected at the point a.

Proof. By theorem B, condition 1 implies condition 2. Consequently it is sufficient to prove the theorem for each of the conditions 2 and 3. The proof of the first case follows the lines of the proof of Lemma 3^i of the paper cited in footnote 7 and will not be reproduced here. For the second case, suppose condition 3 is satisfied. Let $\epsilon > 0$ be arbitrarily given. There exists an $\eta > 0$ such that every complete i-cycle on $\overline{S(a, \epsilon)}$ mod $[M - S(a, \epsilon)]$ bounds on M mod $[M - S(a, \eta)]$. Let G be a neighborhood of a of diameter $< \eta$ and satisfying the conditions $\dim(\bar{G} - G) = \dim_a M - 1$, $p^i(\bar{G} - G) = m$, finite. The proof then proceeds as in the preceding case.

Remark. The preceding theorem is true even in the case $\beta^i(a, M) = \aleph_0$, if the neighborhood G may be chosen so that its diameter is $< \epsilon$ and $< \eta$ and such that every complete i-cycle on $\bar{G} - G$ bounds in $S(a, \epsilon)$.

As a corollary to the preceding theorem, we have the following well known result.

Corollary. *If M is a compact metric continuum and a is a point of M such that* $\mathrm{ord}_a M$ *is finite or* ω, *then M is locally* 0*-connected at a.*

Definition. A domain D of the compact euclidean space R^n is called *uniformly locally i-connected* (u.l.i-c.) if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that every i-cycle in D of diameter $< \delta$ bounds a chain in D of diameter $< \epsilon$.

Theorem 6. *Let K be a closed subset of R^n, D a domain of $R^n - K$ such that the following conditions are satisfied.*

1. *If a is a point of $\bar{D} - D$, then $\beta^{n-2}(a, K) \leqq 1$;*

2. *If K cuts R^n locally at a, one of the local domains is a subset of a domain D_1 of $R^n - K$ distinct from D and such that $D + D_1$ is a subset of a domain complementary to some relative neighborhood of a in K.*

Then D is uniformly locally 0*-connected.*

Proof. Suppose D is not u.l.0-c. There exists a point a of $\bar{D} - D$ and an $\epsilon > 0$ such that, for every $\sigma > 0$, $S(a, \sigma)D$ contains a 0-cycle which fails to bound in $S(a, \epsilon)D$. Let G be a relative neighborhood of a with respect to K which satisfies 2, is contained in $S(a, \epsilon)$ and is such that $p^{n-2}(\bar{G} - G) \leqq 1$. Let $\sigma > 0$ be chosen so that $S(a, \sigma)K$ is contained in G. Let[18] $x_1^0 + x_2^0$ be a 0-cycle of $S(a, \sigma)D$ which fails to bound in $S(a, \epsilon)D$. Let y_1^0 be a point of $S(a, \sigma)D_1$. Then there exist chains L_1^1 and L_2^1 in $S(a, \sigma)$ such that $L_1^1 \rightarrow x_1^0 + y_1^0$, $L_2^1 \rightarrow x_2^0 + y_1^0$ in $R^n - (K + F(a, \epsilon) - G)$. Let $L^1 = L_1^1 + L_2^1$. Then $L^1 \rightarrow x_1^0 + x_2^0$. Since D is connected, there exists a chain L_0^1 in D such that $L_0^1 \rightarrow x_1^0 + x_2^0$ in $R^n - \bar{G}$. Then $L^1 + L_0^1$ is a 1-cycle in $R^n - (\bar{G} - G)$. If

[17] See footnote 1, p. 2.

[18] For simplicity the proofs of the following theorems are stated in terms of mod 2 topology, i.e., the coefficient domain consists of the integers mod 2. They may be modified to hold for any finite coefficient domain. Compare with proofs in R. L. Wilder's paper, *A converse of the Jordan-Brouwer separation theorem in three dimensions*, Transactions of the American Mathematical Society, vol. 32 (1930), p. 635.

$L^1 + L_0^1 \sim 0$ in $R^n - (\bar{G} - G)$, it follows from Theorem C that $x_1^0 + x_2^0 \sim 0$ in $R^n - (K + F(a, \epsilon))$ and, consequently, in $S(a, \epsilon)D$, a contradiction.

From 2 there exists a chain L_3^1 in $R^n - \bar{G}$ such that $L_3^1 \to x_1^0 + y_1^0$. Then $L_1^1 + L_3^1$ is a 1-cycle linking $\bar{G} - G$. For if not, by virtue of Theorem C, $x_1^0 + y_1^0$ would bound in $R^n - K$.

Similarly, $L_0^1 + L_2^1 + L_3^1$ links $\bar{G} - G$. Hence, by 1, $L_1^1 + L_3^1 \sim L_0^1 + L_2^1 + L_3^1$ in $R^n - (\bar{G} - G)$ or $L_0^1 + L_1^1 + L_2^1 \sim 0$. This has been shown to lead to a contradiction.

THEOREM 7. *Let K be a closed subset of R^n, D a domain of $R^n - K$ such that the following conditions are satisfied, where r is any fixed integer,* $1 \leqq r \leqq n - 2$.

1. *If a is any point of $\bar{D} - D$, then* $\beta^{n-r-2}(a, K) = 0$.

2. *If a is any point of $\bar{D} - D$, there exists a relative neighborhood G' of a and a real number $\sigma > 0$ such that any r-cycle in $S(a, \sigma)D$ bounds in the complement of $\bar{G}'$.*

Then D is uniformly locally r-connected.

Proof. Suppose D is not u.l.r-c. There exists a point a of $\bar{D} - D$ and an $\epsilon > 0$ such that, for every $\sigma > 0$, $S(a, \sigma)D$ contains an r-cycle which fails to bound in $S(a, \epsilon)D$. Let G be a relative neighborhood of a contained in $S(a, \epsilon)$ and in the G' of 2, and such that $p^{n-r-2}(\bar{G} - G) = 0$. Let $\sigma > 0$ be chosen so that $S(a, \sigma)K$ is contained in G, and less than the σ of 2. Let γ^r be an r-cycle of $S(a, \sigma)D$. There exists a chain L_1^{r+1} in $S(a, \sigma)$ such that $L_1^{r+1} \to \gamma^r$ in $R^n - (K + F(a, \sigma) - G)$ and a chain L_2^{r+1} in $R^n - \bar{G}'$ such that $L_2^{r+1} \to \gamma^r$ in $R^n - \bar{G}$. Then $L_1^{r+1} + L_2^{r+1}$ is an $(r + 1)$-cycle of $R^n - (\bar{G} - G)$ and, by hypothesis, bounds there. Consequently γ^r bounds in $R^n - (K + F(a, \epsilon))$, a contradiction.

8. **Addition theorems.** In this section we interpolate some general addition theorems which are necessary to the further development of our investigation.

DEFINITION. A compact metric space K is said to be an *irreducible membrane* with respect to a complete $(n - 1)$-cycle γ^{n-1} if $\gamma^{n-1} \sim 0$ on K but on no proper closed subset of K.[8]

DEFINITION. An n-dimensional compact metric space M is said to be an n-dimensional *closed cantorian manifold* if $p^n(M) > 0$ while, if M' is any proper closed subset of M, $p^n(M') = 0$. It is said to be regularly closed if $p^n(M) = 1$.[8]

THEOREM 8. *Let J be an $(n - 1)$-dimensional regularly closed cantorian manifold, K_1 and K_2 two n-dimensional irreducible membranes with respect to the essential complete $(n - 1)$-cycle carried by J such that $K_1K_2 = J$ and $p^n(K_i) = 0$, $i = 1, 2$. Then $K_1 + K_2$ is an n-dimensional regularly closed cantorian manifold.*

Proof. We may suppose $K_1 + K_2$ imbedded in the compact euclidean space R^{2n+1}. It is then sufficient to show that

1. $K_1 + K_2$ is linked by an n-cycle in $R^{2n+1} - (K_1 + K_2)$,
2. No proper subset of $K_1 + K_2$ has this property,
3. The n-cycle in condition 1 is unique.

Proof of 1. We apply Theorem D, setting $K_1 = F'$, $K_2 = F''$, $2n + 1 = m$,

$n - 1 = r$. Then $J = F'F''$, and carries an $(n - 1)$-cycle which bounds on K_1 and on K_2 but not on J, and hence $p^n(R^{2n+1} - (K_1 + K_2)) > 0$.

Proof of 2. We apply Theorem E. Let S be any proper closed subset of $K_1 + K_2$ and set $SK_1 = F'$, $SK_2 = F''$, $2n + 1 = m$, $n - 1 = r$. Then $m - r - 2 = n$, and every n-cycle of $R^{2n+1} - S$ bounds in $R^{2n+1} - SK_1$ and in $R^{2n+1} - SK_2$, since $p^n(SK_i) = p^n(K_i) = 0$, $i = 1, 2$, and, consequently, every such cycle bounds in $R^{2n+1} - S$ unless $SK_1K_2 = SJ$ carries an $(n - 1)$-cycle which fails to bound on SJ but which bounds on SK_i, $i = 1, 2$. In order that SJ carry a non-bounding $(n - 1)$-cycle, it is necessary that $SJ = J$. In order that such a cycle bound on SK_i, it is necessary that $SK_i = K_i$. Since S is a proper subset of $K_1 + K_2$, these conditions cannot both be satisfied, and the proof that $K_1 + K_2$ is a closed cantorian manifold is complete. That $K_1 + K_2$ is regularly closed follows from an addition theorem due to Mayer.[19]

In some cases in which the conditions $p^n(K_i) = 0$ are not known to be satisfied, the following corollary is useful.

Corollary. *Let J, K_1 and K_2 satisfy the hypotheses of the preceding theorem except that $p^n(K_i)$ is not required to be zero. Then $K_1 + K_2$ is the irreducible carrier of an essential complete n-cycle.*

Proof. This proof is essentially the same as that of the theorem. It is only necessary to make use of the fact that the linking cycle given by Theorem D bounds in the complement of K_i and hence, part 1 being as before, in part 2 this cycle bounds in the complement of SK_i. This leads as before to a contradiction unless $S = K_1 + K_2$.

The preceding theorem may be generalized as follows.

Theorem 9. *Let J be the carrier of a complete $(n - 1)$-cycle which fails to bound on J, and let K_1 and K_2 be two n-dimensional irreducible membranes with respect to this cycle such that $K_1K_2 = J$, $p^n(K_i) = 0$, $i = 1, 2$. Furthermore, suppose that K_1 and K_2 are irreducible membranes with respect to any complete $(n - 1)$-cycle carried by J which fails to bound on J but which bounds on K_1 and on K_2. Then $K_1 + K_2$ is an n-dimensional closed cantorian manifold. Also, $p^n(K_1 + K_2)$ is the number of $(n - 1)$-cycles of the type described.*

Proof. The proof follows the same lines as that of the preceding theorem. That of part 1 may be used as it stands. In the proof of part 2, there is the alternative that SJ may contain an $(n - 1)$-cycle which fails to bound on SJ but which bounds on J, on SK_1 and on SK_2. In this case the argument of part 1 shows that $J + SK_1$, which is contained in K_1, carries a non-bounding n-cycle. This contradicts the assumption that K_1 is n-dimensional and that $p^n(K_1) = 0$. The last part of the theorem follows as before.[18]

The corollary of Theorem 8 can be extended in the case of the preceding theorem.

[19] Monatshefte für Mathematik und Physik, vol. 36 (1929), p. 40. See also Whyburn, *Cyclic elements of higher orders*, American Journal of Mathematics, vol. 36 (1934), p. 136, footnote.

Combining the above theorem with one due to Alexandroff,[20] we get

THEOREM 10. *The necessary and sufficient condition that the compact metric space M be an n-dimensional closed cantorian manifold is as follows.*

1. $M = K_1 + K_2$ *with* K_i, $i = 1, 2$, *n-dimensional compact metric spaces such that* $p^n(K_i) = 0$.

2. *Every complete* $(n-1)$*-cycle carried by* K_1K_2 *which fails to bound on* K_1K_2 *but which bounds on* K_1 *and on* K_2 *has these sets as irreducible membranes.*

3. *At least one such cycle as described in* 2 *exists.*

The lower dimensional connectivities of a closed cantorian manifold considered as the sum of two irreducible membranes may be found by applying the Mayer addition theorem.[18] The following special case may also be proved by the use of Theorem C:

THEOREM 11. *If, in addition to the hypotheses of Theorem* 9, $p^{n-r-1}(K_1K_2) = 0$ *and* $p^{n-r}(K_i) = 0$, $i = 1, 2$, *then* $p^{n-r}(K_1 + K_2) = 0$.

Proof. To apply Theorem C, let $K_1 = A$, $K_2 = B$, $2n + 1 = m$, $n + r = k$. Then $p^{n+r+1}(R^{2n+1} - K_1K_2) = p^{n-r-1}(K_1K_2) = 0$, $p^{n+r}(R^{2n+1} - K_i) = p^{n-r}(K_i) = 0$, and the statement follows.

9. Application of addition theorems. Using Theorem F it is possible to obtain the following extremely useful result.

THEOREM 12. *Let M be an n-dimensional compact metric space which is the irreducible carrier of a complete n-cycle which fails to bound on M. If a is a point of M such that* $\beta^{n-1}(a, M)$ *is finite or* ω, *then M is locally* 0*-connected at a.*

Proof. This follows from the fact that arbitrarily small neighborhoods of a may be chosen whose boundaries have finite $(n-1)$-dimensional Betti numbers, and consequently, by Theorem F, separate M into a finite number of components. That component of such a separation which contains the point a has a diameter at most equal to that of the neighborhood whose boundary determines the separation and is itself a connected neighborhood of a. Consequently a has arbitrarily small connected neighborhoods and is a point of local 0-connectedness of M.

The question now arises as to whether or not the local condition in the hypothesis of the preceding theorem is sufficient to insure the same conclusion for other classes of compact metric spaces. The following theorem answers this in the affirmative, making use of the addition theorems already developed.

THEOREM 13. *Let M be an n-dimensional compact metric space with* $p^n(M) = 0$, *J a locally* 0*-connected closed subset of M which carries a complete* $(n-1)$*-cycle which fails to bound on J but with respect to which M is an irreducible membrane, and such that every complete* $(n-1)$*-cycle carried by J which fails to bound on J but bounds on M has M for an irreducible membrane. Suppose further that if a is any point of M,* $\beta^{n-1}(a, M)$ *is finite or* ω. *Then M is locally* 0*-connected at each of its points.*

Proof. M is locally 0-connected at all points of $M - J$. Let M' be a set

[20] See footnote 9, p. 186.

homeomorphic to M and so situated that $MM' = J$. Then by Theorem 9 $M + M'$ is a closed cantorian manifold, at every point a of which, except possibly those of J, $\beta^{n-1}(a, M + M')$ is finite or ω. From the previous theorem it follows that $M + M'$ is locally 0-connected at each such point and the same is true of M itself.

The space M is locally 0-connected at each point of the set J. Suppose that a is a point of J at which M is not locally 0-connected. There exists an $\epsilon > 0$ such that any neighborhood of a of diameter less than ϵ has an infinite number of components. Let $\delta > 0$ be chosen corresponding to ϵ with respect to the local 0-connectedness of J at the point a. Let G be a neighborhood of a contained in $S(a, \delta)$ such that $\dim(\bar{G} - G) = n - 1$ and $p^{n-1}(\bar{G} - G) = m$, finite. There exists an infinite sequence, (g_i), of components of G. At least $m + 1$ of these have no limit points (and hence no points) on GJ. For, if all but a finite number had such points, they might be added to the component of $S(a, \epsilon)J$ determined by a, and it would follow that any two points of M sufficiently near to a would belong to a connected subset of $S(a, \epsilon)$ and M would be locally 0-connected at a. Now let g be one of these $m + 1$ components. All of its limit points in G belong to it, and none of its limit points in $\bar{G} - G$ belong to J. Moreover, no point of g is a limit point of $M - g$, since such a point would lie in $G - J$ and hence in $M - J$ and be a point of non-local 0-connectedness of M in $M - J$. Consequently $(\bar{G} - G)\bar{g}$ separates g from M. If the set M' of the preceding paragraph is again added to M, it follows that g is separated by the same set from $M + M'$ since, having no limit points on $J = MM'$, g can have none on M'. By Theorem F it follows that[21] $p^{n-1}(\bar{G} - G) \geqq m + 1$.

The condition that J be itself locally 0-connected is necessary, as is shown by the following example. Let M be the compact plane set whose boundary, taken as J, consists of the following three parts: (1) the curve $y = \sin 1/x$, $0 < x \leqq 1/\pi$, (2) the segment $x = 0$, $-1 \leqq y \leqq 1$, (3) the arc $(x - 1/2\pi)^2 + (y - \frac{1}{2})^2 = 1/4\pi^2$, $y \leqq -\frac{1}{2}$.

10. Characterization of the 2-manifold.

THEOREM 14. *Let M be an n-dimensional compact metric space such that $p^n(M) = m > 0$, while if M' is any proper closed subset of M, $p^n(M') < m$. Let a be a point of M. There exists a positive integer $k \leqq m$ such that, if G is any sufficiently small neighborhood of a, $p^n(M - G) = m - k$, and $\beta^{n-1}(a, M) \geqq k$.*

Proof. The existence of the number k follows from the fact that, as the diameter of G decreases, $p^n(M - G)$ increases, or remains constant, but never exceeds the value $m - 1$. That $\beta^{n-1}(a, M) \geqq k$ follows from the Mayer addition theorem,[18] since $\bar{G} - G$ must carry at least k complete $(n - 1)$-cycles which fail to bound on $\bar{G} - G$.

COROLLARY. *Let M be an n-dimensional closed cantorian manifold such that, for every point a of M, $\beta^{n-1}(a, M) \leqq 1$. Then M is regularly closed and locally 0-connected.*

[21] See footnote 9, p. 153.

Proof. If $p^n(M) = m$, for every point a of M, $k = m$. But this implies, by the preceding theorem, that $\beta^{n-1}(a, M) \geqq m$, and consequently, $m = 1$. The local 0-connectedness of M follows from Theorem 12.

COROLLARY. *Let M be an n-dimensional closed cantorian manifold imbedded in R^{n+1} and such that, for every point a of M, $\beta^{n-1}(a, M) \leqq 1$. Then M separates R^{n+1} into exactly two uniformly locally 0-connected complementary domains, of which it is the common boundary.*

Proof. It follows from the preceding corollary and Theorem 6.

COROLLARY. *Let M be a 2-dimensional locally 1-connected closed cantorian manifold imbedded in R^3 and such that, for each point a of M, $\beta^1(a, M) \leqq 1$. Then M is a 2-dimensional combinatorial manifold.*

Proof. This follows from the preceding corollary and a theorem due to Wilder.[22]

The following theorem shows, as might be expected, that the restriction in the preceding corollary that M be imbedded in R^3 is unnecessary.

PRINCIPAL THEOREM A. *Let M be a 2-dimensional closed cantorian manifold, such that, if a is any point of M, $\beta^1(a, M) \leqq 1$. Then M is a 2-dimensional combinatorial manifold.*

Proof. As immediate consequences of the hypotheses and of Theorem 12, it follows that M is locally 0-connected and that, for every point a of M, $\beta^1(a, M) = 1$. By Theorem G it is sufficient to show the existence of a real number $\delta > 0$ such that every simple closed curve of M of diameter $< \delta$ cuts M. Assuming that this is false, there exists a point a of M such that, for every real number $\delta > 0$, $S(a, \delta)$ contains a simple closed curve which fails to cut M.

Let $\epsilon > 0$ be arbitrarily chosen. Then, since M is locally 1-connected, $\delta > 0$ may be chosen in such a manner that every complete 1-cycle carried by $S(a, \delta)$ bounds on $S(a, \epsilon)$. By hypothesis, $S(a, \delta)$ contains a simple closed curve J which fails to cut M. The essential complete 1-cycle carried by J bounds in $S(a, \epsilon)$ and there exists a subset K of $S(a, \epsilon)$ which is an irreducible membrane with respect to this cycle.

Let p' be a point of $K - J$, r a point of $M - K$. Since $M - J$ is a connected open subset of the Peano continuum M, there is an arc α', with end points p' and r, in $M - J$. Let p be the first point of K, consequently a point of $K - J$, on α' in the direction from r to p'. Let α denote the subarc $\widehat{pr}$ of α'.

Let $\eta > 0$ be so chosen that $\eta < \frac{1}{2}\rho(p, J + r)$. Let G be a neighborhood of p such that (1) $G \subset S(p, \eta)$, (2) $\dim(\bar{G} - G) = 1$, (3) $p^1(\bar{G} - G) = 1$. Let $\gamma^2 = (\gamma_1^2, \cdots, \gamma_n^2, \cdots)$ be an essential complete 2-cycle carried by M, γ_n^2 being an ϵ_n-cycle with $\lim_{n \to \infty} \epsilon_n = 0$. By an ϵ_n-transformation of those vertices of γ_n^2 which are within a distance ϵ_n of $\bar{G} - G$, we may insure that each cell of γ_n^2 either has all of its vertices on $\bar{G}$, or none of its vertices on G. Let l_n^2 denote the subcomplex of γ_n^2 composed of all cells of the former class. The boundary of l_n^2 is then an ϵ_n-cycle i_n^1 on $\bar{G} - G$. By making a proper choice of a subsequence of the cycles i_n^1, it is possible to obtain a complete 1-cycle $i^1 = (i_1^1, \cdots, i_n^1, \cdots)$.

[22] See footnote 10, p. 306.

Since M is an irreducible carrier of the complete 2-cycle γ_n^2, it follows that $\bar{G}$ is an irreducible membrane with respect to the complete cycle i^1.

Let $j^1 = (j_1^1, \cdots, j_n^1, \cdots)$ be an essential complete 1-cycle on J, where j_n^1 is an ϵ_n-cycle, and let k_n^2 be an ϵ_n-chain realizing the homology $j_n^1 \underset{\epsilon_n}{\sim} 0$ irreducibly on K. As before, by an ϵ_n-transformation of the vertices of k_n^2, it is possible to insure that each cell of k_n^2 either has all of its vertices on $\bar{G}$ or none of its vertices on G. Let $\bar{k}_n^2$ denote the subcomplex of k_n^2 consisting of all cells of the former class. The boundary of k_n^2 is then an ϵ_n-cycle $\bar{i}_n^1$ on $\bar{G} - G$. We again suppose a proper subsequence of the cycles $\bar{i}_n^1$ to be chosen in such a way as to form a complete 1-cycle $\bar{i}^1 = (\bar{i}_1^1, \cdots, \bar{i}_n^1, \cdots)$.

From (2) it follows that there exists an ϵ_n-complex, m_n^2, on $\bar{G} - G$ such that $m_n^2 \to i_n^1 + \bar{i}_n^1$ on $\bar{G} - G$. Moreover, $\gamma_n^2 + \bar{k}_n^2 \to \bar{i}_n^1$ on M, $l_n^2 + m_n^2 \to \bar{i}_n^1$ on $\bar{G}$. Since $\alpha - p$ is on the carrier of γ_n^2 but not on the carrier of $\bar{k}_n^2$, it is a part of the carrier of $\gamma_n^2 + \bar{k}_n^2$. Then $\gamma_n^2 + \bar{k}_n^2 + l_n^2 + m_n^2 \to 0$, and this cycle is carried by $M - (\alpha - p)G$, a proper closed subset of M. Since this cycle differs from the non-bounding complete 2-cycle γ^2 only in a small neighborhood of a, it is also a non-bounding complete 2-cycle. This, however, contradicts the fact that M is a 2-dimensional closed cantorian manifold. This proves the theorem.

Corollary. *In the hypothesis of the preceding theorem the condition that M be locally* 1*-connected may be replaced by any one of the conditions* 1, 2 *and* 3 *of Theorem* 5 *(for $i = 1$).*

Proof. Since $\beta^1(a, M)$ is required to be not greater than 1 for every point a of M, the hypothesis of Theorem 5 is satisfied and M is locally 1-connected at each of its points.

Principal Theorem A can be stated in several ways. The following statement brings out some points of interest.

Theorem 15. *Let M be a compact metric space satisfying the following conditions:*

1. dim $M = 2$,
2. $p^2(M) > 0$, *but, if M' is any proper closed subset of M, $p^2(M') = 0$,*
3. $p^1(M)$ *is finite,*
4. *if a is any point of M, $\beta^1(a, M) \leqq 1$.*

Then M is a 2*-dimensional combinatorial manifold.*

Thus we begin with the point set notion of a 2-dimensional compact metric space, and by subjecting it to certain combinatorial conditions obtain the class of 2-dimensional combinatorial manifolds. Moreover, the only local restriction, except for the dimension, is that supplied by the number $\beta^1(a, M)$. It is to be noted that if we wish to characterize any particular type of manifold, such as the sphere, we need only require $p^1(M)$ to have some particular value, in this case zero, and, in some cases, also require orientability or non-orientability.[23]

11. Characterization of the closed 2-cell.

Principal Theorem B. *Let M be a* 2*-dimensional compact metric space with $p^2(M) = p^1(M) = 0$, J a simple closed curve contained in M and such that*

[23] Veblen, *Analysis Situs*, 2nd ed., p. 50.

M is an irreducible membrane with respect to some essential complete 1*-cycle carried by J. Suppose also that, if a is a point of M, $\beta^1(a, M) \leqq 1$, while, in particular, if a is a point of J, $\beta^1(a, M) = 0$. Then M is a closed* 2*-cell.*

Proof. The condition that M be locally 1-connected, which was necessary in the hypothesis of Principal Theorem A, is here replaced by the stronger and certainly necessary condition $p^1(M) = 0$.

Since M is an irreducible membrane with respect to an essential complete 1-cycle carried by J, it is an irreducible membrane with respect to any such cycle, since all of them are homologous on J.

Let C be a 2-cell bounded by J and such that $MC = J$. From Theorem 8 it follows that $M + C$ is a 2-dimensional closed cantorian manifold, while from Theorem 11 it follows that $p^1(M + C) = 0$. Moreover, $M + C$ is evidently locally 1-connected, and, if a is any point of $M + C - J$, $\beta^1(a, M + C) = 1$. If this equality can be shown to hold for each point of J, it will follow from Principal Theorem A that $M + C$ is a 2-sphere and, consequently, that M is a closed 2-cell.

By Theorem 13, M is locally 0-connected at all points so that, if a is a point of J and $\epsilon > 0$, there exists a connected neighborhood G of a such that $\delta(G) < \epsilon$, $\dim(\bar{G} - G) = 1$ and $p^1(\bar{G} - G) = 0$. (All neighborhoods are with respect to M.)

It is first necessary to show that no point of J is a local cut point of M. To do this, suppose that a is a point of J which is a local cut point of M. Let G_1 be a connected neighborhood of a such that $\dim(\bar{G}_1 - G_1) = 1$, $p^1(\bar{G}_1 - G_1) = 0$ and $G_1 - a$ is not connected. Let G_2 be a neighborhood of a contained in G_1, having the same properties and also being sufficiently small so that JG_2 is contained in the component of JG_1 determined by a. It follows that at most two components of $G_1 - a$ have points in common with the set $(G_2 - a)J$. Also, if a component of $G_1 - a$ has, as its only limit point on G_2J, the point a, the portion of this component in G_2 is separated from the closed cantorian manifold $M + C$ by a set consisting of the point a and a subset of the boundary of G_2 which has no point on J. This subset must then have a positive first Betti number.[24] This contradicts the hypothesis that $\dim(\bar{G}_2 - G_2) = 1$ and $p^1(\bar{G}_2 - G_2) = 0$. Since G_1 is connected, every component of $G_1 - a$ must either contain points of $(G_2 - a)J$ or have a as its only limit point. The preceding analysis therefore shows that $G_2 - a$ has exactly two components determined by the two arcs of $G_2J - a$. The closures of these components will be denoted by A_1 and A_2.

As the essential complete 1-cycle on J, we may take a cycle $\gamma = (\gamma_1, \cdots, \gamma_n, \cdots)$ where γ_n is an ϵ_n-cycle, $\epsilon_n \to 0$, whose vertices are arranged in a definite cyclic order on J and include the point a. Let C_n be an ϵ_n-chain on M bounded by γ_n. Make an ϵ_n-deformation of the vertices of C_n so that any cell of C_n either has all its vertices on A_1 or none of its vertices on A_1G_1. This deformation may be carried out so as not to affect vertices of γ_n near a. Let $\bar{C}_n$ be the subchain of C_n consisting of all cells of the latter whose vertices are on A_1. The

[24] See footnote 21.

boundary of $\bar{C}_n$ consists of a cycle which is the sum of two chains, one on $\bar{G}_1 - G_1$, the other on J. The carrier of the latter must contain a subarc of J ending at a, and a carries a boundary vertex of this chain. However, a cannot carry a boundary vertex of the other chain since this chain is at a positive distance from a. Consequently the sum of the two chains cannot be a cycle and we reach a contradiction. Therefore a is not a local cut point of M.

Since this is true, it follows that, if $\sigma > 0$, $S(a, \sigma)$ contains an arc α in $S(a, \sigma) - a$ joining two points p and q of the component of $JS(a, \epsilon)$ determined by a which are separated in this component by the point a. Let $\epsilon > 0$ be chosen arbitrarily and let the above σ be so chosen that, if β is the arc $\widehat{paq}$, any complete 1-cycle carried by $\alpha + \beta$ bounds in $S(a, \epsilon)$. That this is possible follows from the local 1-connectedness of M. Now suppose $\alpha + \beta$ does not separate M. It is possible to repeat, word for word, the steps in the proof of Theorem 14, replacing J in that theorem by $\alpha + \beta$.

So far we have $(M + C) - (\alpha + \beta)$ separated, (C being a 2-cell bounded by J, $MC = J$), one of the components, say M_1, being such that its closure is a subset of M of diameter less than ϵ, which is an irreducible membrane with respect to an (any) essential complete 1-cycle on $\alpha + \beta$. Since $\alpha + \beta$ separates M_1 from the remainder of $M + C$, which contains the arc $J - \beta$, the only limit points of M_1 on J are points of β. By Theorem F, $(M + C) - (J + \alpha)$ has, from the connectivity of J, at most three components. One of these is C, another M_1, and the third, M_2, is the remainder of M.

Let (γ_n), (γ_n^1), (γ_n^2) be essential complete 1-cycles on J, $\alpha + \beta$, and $(J - \beta) + \alpha = J_1$, respectively, and such that $\gamma_n = \gamma_n^1 + \gamma_n^2$.

γ_n bounds irreducibly on M, γ_n^1 bounds irreducibly on $\bar{M}_1$. Consequently, their difference, γ_n^2, bounds on M. Let K be an irreducible membrane with respect to γ_n^2 contained in M. Since M is an irreducible membrane with respect to γ_n, it follows that K must contain $\bar{M}_2$, since K is closed and since the carrier of the homology $\gamma_n^1 \sim 0$ is $\bar{M}_1$.

If α does not separate M, then $\bar{M}_2$ contains interior points of the arc β. Let r be one such point. Let C' be a 2-cell bounded by J_1 and such that $MC' = J_1$. By Theorem 8, $K + C'$ is a closed cantorian manifold, the hypothesis $p^2(K) = 0$ of the theorem being satisfied, since M, which contains K, is 2-dimensional and $p^2(M) = 0$.

Let G be any neighborhood of r in $M + C'$ so small that $\bar{G}C' = 0$, and such that $\dim(\bar{G} - G) = 1$. Then G is a neighborhood of r in M and KG is a neighborhood of r in $K + C'$. Consequently, $p^1(\overline{KG} - KG) > 0$ and $p^1(\bar{G} - G) > 0$. But this contradicts the hypothesis that $\beta^1(r, M) = 0$. Therefore $M - \alpha = M_1 + (\beta - p - q) + M_2$.

It follows that if a is a point of J, there exist arbitrarily small neighborhoods of a in M whose boundaries are arcs having both end points on J. Each of these may be extended to a neighborhood of a in $M + C$ whose boundary is a simple closed curve. Consequently $\beta^i(a, M + C) \leqq 1$, and the theorem is proved.

It is interesting to note that Alexandroff[25] raised the question as to whether or not a locally 0- and 1-connected set M which is an irreducible membrane with respect to a simple closed curve and satisfies the conditions $p^1(M) = p^2(M) = 0$ is necessarily a closed 2-cell. This is answered in the negative by the example shown below.

The same method may be used to give characterizations of sets obtained by omitting a finite number of open 2-cells from 2-manifolds. In general, J will be replaced by a finite number of simple closed curves and $p^1(M)$ will be required to have some non-zero finite value. The local conditions remaining as in Principal

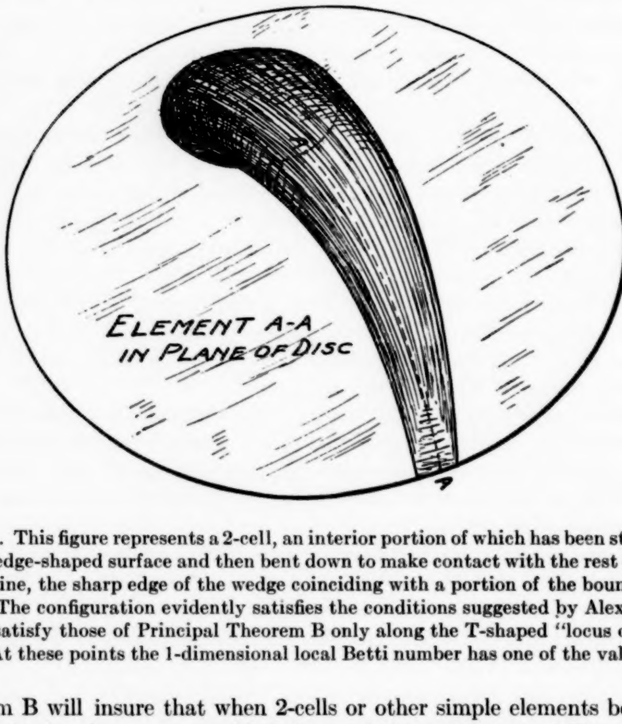


FIG. 1. This figure represents a 2-cell, an interior portion of which has been stretched out into a wedge-shaped surface and then bent down to make contact with the rest of the 2-cell along a line, the sharp edge of the wedge coinciding with a portion of the boundary of the 2-cell. The configuration evidently satisfies the conditions suggested by Alexandroff and fails to satisfy those of Principal Theorem B only along the T-shaped "locus of singularities". At these points the 1-dimensional local Betti number has one of the values 2 and 3.

Theorem B will insure that when 2-cells or other simple elements bounded by the simple closed curves are added the conditions of Principal Theorem A will be satisfied. In some cases it may be necessary to make some hypothesis concerning orientability. As an example, we give the following

THEOREM 16. *Let M be a 2-dimensional compact metric space with $p^2(M) = 0$, $p^1(M) = 1$, J a simple closed curve contained in M and such that M is an irreducible membrane with respect to an essential complete 1-cycle carried by J. Suppose also that, if a is a point of M, $\beta^1(a, M) \leqq 1$, and, if a is a point of J, $\beta^1(a, M) = 0$. Then M is homeomorphic to a Moebius strip.*

[25] See footnote 9, p. 181.

Proof. Let C be a 2-cell bounded by J such that $MC = J$. Just as in the previous theorem it can be shown that $M + C$ is a combinatorial manifold, while, from the Mayer addition theorem, $p^2(M + C) = 1$. Consequently, $M + C$ is a projective plane, and M is a Moebius strip.

12. Characterization of the open 2-cell.

The following theorem is frequently useful.

THEOREM 17. *Let M be an n-dimensional locally compact (or compact) metric continuum such that, if a is any point of M, $\beta^{n-1}(a, M) = 1$, and, in fact, if $\epsilon > 0$, a neighborhood G of a exists such that $\delta(G) < \epsilon$, $\dim(\bar{G} - G) = n - 1$, $p^{n-1}(\bar{G} - G) = 1$ and $\bar{G}$ is an irreducible membrane with respect to a non-bounding complete $(n - 1)$-cycle of $\bar{G} - G$. Then if J is any set which carries a complete $(n - 1)$-cycle which bounds on a set K such that $K - J$ and $M - (K + J)$ are non-vacuous, $M - J$ is not connected.*

Remark. *It is sufficient to assume that the local condition applies only to points of $K - J$.*

Proof. Suppose $M - J$ is connected. Let r be a point of $M - K$. Since J does not cut M, the component of $M - K$ containing r has a limit point p on $K - J$. Take $\epsilon < \rho(p, J)$ and let G be a neighborhood of p satisfying the hypothesis of the theorem. Then $\bar{G}$ is an irreducible membrane with respect to the cycle on $\bar{G} - G$. However, $\bar{G}K$ is also an irreducible membrane with respect to this cycle, since K is an irreducible membrane with respect to the cycle on J. But $\bar{G}K$ is a proper subset of $\bar{G}$, since it contains no points of the component of $M - K$ containing r, while G does. This contradiction proves the theorem.

Making use of this theorem we obtain a characterization of sets all points of which have 2-cell neighborhoods, as follows.

PRINCIPAL THEOREM C. *Let M be a 2-dimensional locally compact metric continuum such that, if a is any point of M, $\beta^1(a, M) = 1$ and, in fact, such that for every $\epsilon > 0$ there exists a (compact) neighborhood G of a of diameter $< \epsilon$ such that* $\dim(\bar{G} - G) = 1$, $p^1(\bar{G} - G) = 1$, $p^1(\bar{G})$ *is finite and $\bar{G}$ is an irreducible membrane with respect to a complete cycle on $\bar{G} - G$. Then every point of M has a 2-cell neighborhood.*

Proof. Let a be a point of M, G_1 a neighborhood of a of the type described. By Theorem 5, $\bar{G}_1$ is locally 1-connected.

Let G be a neighborhood of a of the type described and so small that every complete 1-cycle on G bounds on G_1. Then Theorem 17 applies for every J contained in G, i.e., $\bar{G}_1 - J$ is not connected. We will show shortly that M is locally 0-connected. From this it follows that $G - J$ is not connected, and, in particular, that any simple closed curve of G separates G.

From the hypotheses on $\bar{G} - G$ it follows that this set contains a regularly closed 1-dimensional cantorian manifold C.[26] Let $\bar{G}'$ be an auxiliary set homeomorphic with $\bar{G}$ and such that $\overline{GG'} = C$. By the corollary to Theorem 8,

[26] Wilder, *Point sets in three and higher dimensions*, Bulletin of the American Mathematical Society, vol. 38 (1932), pp. 649–692; see bottom p. 681.

$\bar{G} + \bar{G}'$ is the irreducible carrier of an essential complete 2-cycle. By Theorems 12 and F it follows that G is locally 0-connected and is cut by no arc of G.

Since G is locally 0-connected, locally compact, and cut by every simple closed curve but by no arc, it follows from Theorem H that G is a cylinder-tree, and, consequently, that every point of M has a 2-cell neighborhood.

It is evident that, if every point of M has a 2-cell neighborhood, then M satisfies all the hypotheses of the theorem. These conditions are then necessary and sufficient. In place of the condition "$p^1(\bar{G})$ is finite" the equivalent hypothesis "M is locally 1-connected" might be used.

Principal Theorem D. *Let M be a 2-dimensional locally compact, non-compact, metric continuum such that every complete 1-cycle carried by M bounds on a compact set on M and such that, if a is any point of M and $\epsilon > 0$ any real number, there exists a (compact) neighborhood G of a of diameter less than ϵ such that* $\dim(\bar{G} - G) = 1$, $p^1(\bar{G} - G) = 1$, *and $\bar{G}$ is an irreducible membrane with respect to a complete 1-cycle on $\bar{G} - G$. Then M is an open 2-cell.*

Proof. By Theorems 5 and 18 any point of M has a 2-cell neighborhood and, by Theorem 17, every simple closed curve on M cuts M. By Theorem I it remains to show that, if J is a simple closed curve in M, $M - J$ has just two components. To do this, we note that, since M is locally 0-connected, every component of $M - J$ has every point of J as a limit point. For, if not, there is a component C of $M - J$ and a point p of J which is a limit point of C and an end point of an arc of J, no interior point of which is a limit point of C. Let U be a 2-cell neighborhood of p. Since U is a 2-cell, exactly two components of $U - UJ$ have p as a limit point and each of these has as limit points all points of an arc of J to which p is interior. This contradiction proves the statement in question. Now suppose that $M - J$ has three components. Each has the point p of J as limit point, but we have just seen that in any 2-cell neighborhood of p there are only two such components.

Principal Theorem C can also be used to give a characterization of cylinder-trees, i.e., subsets of the 2-sphere which are complementary to closed, totally disconnected sets. All that is necessary is to add the condition that M is imbeddable in R^2. This is conveniently done by means of Theorem J.

Principal Theorem E. *Let M be as described in Theorem* 18 *and, in addition, contain no primitive skew curve. Then M is a cylinder-tree.*

13. Some properties of Alexandroff's local Betti numbers.[1]

Theorem 18. *Let K be a closed subset of R^n, D a domain of $R^n - K$ such that, if a is any point of $\bar{D} - D$,* $p^{n-i-1}(a, K) = 0$. *Then D is uniformly locally i-connected.*

Proof. If D is not u.l.i-c., there exists a number $\epsilon > 0$ and a point a of $\bar{D} - D$ such that, if $\sigma < \epsilon$, $S(a, \sigma) - K$ contains an i-cycle which does not bound in $S(a, \epsilon) - K$. However, since $p^{n-i-1}(a, K) = q^i(a, R^n - K) = 0$, an $\epsilon' < \epsilon$ and a $\sigma' < \epsilon'$ and $< \sigma$ exist such that every i-cycle in $S(a, \sigma') - K$ bounds in $S(a, \epsilon') - K$, and, consequently, in $S(a, \epsilon) - K$.

As a converse theorem we have

THEOREM 19. *Let K be a closed subset of R^n, a a point of K such that only a finite number of domains of $R^n - K$ have a as a boundary point and each such domain is uniformly locally i-connected, $i \neq 0$. Then $p^{n-i-1}(a, K) = 0$.*

Proof. Suppose $p^{n-i-1}(a, K) \neq 0$. Then $q^i(a, R^n - K) \neq 0$ and for every $\epsilon > 0$ there exists a $\sigma > 0$ such that, if $\sigma' < \sigma$, $S(a, \sigma') - K$ contains an i-cycle which fails to bound in $S(a, \epsilon) - K$. This cycle may be taken as irreducible. Since $i \neq 0$ we may suppose it is a connected set and therefore contained in some complementary domain of $R^n - K$. Since there is, by hypothesis, only a finite number of domains having a as a boundary point one of these contains arbitrarily small i-cycles of the above type and hence is not locally i-connected.

For uniform local 0-connectedness we have the stronger result, supplementing Theorem 18,

THEOREM 20. *Let K be a closed subset of R^n, D a domain of $R^n - K$ such that if a is any point of $\bar{D} - D$, it is a boundary point of a finite number, exactly k_a, of domains of $R^n - K$ and $p^{n-1}(a, K) = k_a - 1$. Then D is uniformly locally 0-connected.*

Proof. If D is not u.l.0-c., there exists a point a of $\bar{D} - D$ and a number $\epsilon > 0$ such that, if $\sigma < \epsilon$, $D[S(a, \sigma) - K]$ contains a 0-cycle which does not bound in $D[S(a, \epsilon) - K]$. However, there exists an $\epsilon' < \epsilon$ and a $\sigma' < \epsilon'$ and $< \sigma$ such that there are exactly $k_a - 1$ 0-cycles in $S(a, \sigma') - K$ which are independent in $S(a, \epsilon') - K$. But these must consist of pairs of points in different domains and, consequently, any 0-cycle of $D[S(a, \sigma') - K]$ bounds in $S(a, \epsilon') - K$ and consequently in $D[S(a, \epsilon) - K]$.

As a converse to this theorem we have

THEOREM 21. *Let K be a closed subset of R^n, a a point of K which is a boundary point of a finite number, k_a, of domains of $R^n - K$, each domain being uniformly locally 0-connected. Then $p^{n-1}(a, K) = k_a - 1$.*

Proof. For every $\epsilon > 0$ there exists a $\sigma > 0$ such that every 0-cycle in $S(a, \sigma)$ which is contained in a single domain of $R^n - K$ bounds in $S(a, \epsilon)$. Hence $p^{n-1}(a, K) \leqq k_a - 1$, and the equality is an obvious conclusion.

From these theorems we obtain a characterization of those of R. L. Wilder's generalized closed $(n - 1)$-manifolds[27] which can be imbedded in R^n.

THEOREM 22. *The necessary and sufficient condition that a closed set M in R^n be a generalized closed $(n - 1)$-manifold is*

1. $p^{n-1}(M) = 1$, *while, if M' is any proper closed subset of M, $p^{n-1}(M') = 0$.*
2. *If a is any point of M, $p^{n-i-1}(a, M) = 0$, $1 \leqq i \leqq n - 2$, $p^{n-1}(a, M) = 1$.*

DEFINITION. Let M be a closed subset of R^n, N_1 a neighborhood in R^n of the point a of M. A cycle γ^r of $N_1 - M$ is said to be *irreducibly linked with M at a* if γ^r does not bound in $N_1 - M$ but bounds in $N_1 - (M - G)$, where G is an arbitrarily small relative neighborhood of a.

[27] Wilder, *Generalized closed manifolds in n-space*, Annals of Mathematics, vol. 35 (1934), pp. 876–903.

With respect to the β's we have the following

THEOREM 23. *Let M be a closed subset of R^n and a a point of M at which M is irreducibly linked by m independent (in $N_1 - M$) r-cycles $\gamma_1^r, \gamma_2^r, \cdots, \gamma_m^r$. Then $\beta^{n-r-2}(a, M) \geqq m$.*

Proof. Replace M by the boundary of N_1 together with the points of M interior to N_1. From now on this set will be denoted by M. Let G be any relative (to M) neighborhood of a interior to N_1 and such that the distance of any point of G from a is less than the distance of the γ_i^r from a. Let M' be the relative boundary of G, $G' = M - (G + M')$. Each of the γ_i^r bounds in $R^n - (G' + M')$ by hypothesis and also in $R^n - (G + M')$, since $G + M'$ is contained in a sphere which excludes all the γ_i^r. Therefore each bounds a chain K_i^{r+1} in $R^n - (G + M')$ and a chain $\bar{K}_i^{r+1}$ in $R^n - (G' + M')$. Then $K_i^{r+1} + \bar{K}_i^{r+1}$ is a cycle in $R^n - M'$. If any linear combination of these cycles bounds in $R^n - M'$, the corresponding linear combination of the γ_i^r's bounds in $R^n - M$, by Theorem C, a contradiction which proves the theorem.

The following theorem gives a relation between the β's and Alexandroff's local Betti numbers.[28]

THEOREM 24. *If M has no $(n - r - 1)$-dimensional condensation at a and $p_a^{n-r-1}(M)$ is finite, then $\beta^{n-r-2}(a, M) \geqq p_a^{n-r-1}(M)$. If $p_a^{n-r-1}(M)$ is infinite, $\beta^{n-r-2}(a, M) = \omega$ or $\aleph_0$, depending on whether or not the base determining $p_a^{n-r-1}(M)$ can be so chosen that there exists a $\sigma > 0$ such that all cycles of the base have points in $M - S(a, \sigma)$.*

Proof. Suppose that M has no $(n - r - 1)$-dimensional condensation at a and $p_a^{n-r-1}(M) = m$, finite. Then $p_a^r(R^n - M) = m$. Let $\gamma_i^r = (\gamma_{i1}, \cdots, \gamma_{ik}, \cdots)$, $(i = 1, 2, \cdots, m)$, be a base at a (in $R^n - M$). Let $\epsilon > 0$ be any real number. The sequences γ_i^r may be assumed to be such that every pair of cycles of the sequence γ_i^r are homologous in $S(a, \epsilon) - M$. Take σ so small that some cycle of each sequence lies in $S(a, \epsilon) - \overline{S(a, \sigma)}$. For simplicity of notation this may be assumed to be γ_{i1}. Let G be a neighborhood of a in $S(a, \sigma)$. Let $\sigma' > 0$ be chosen so that $S(a, \sigma')M$ is contained in G. Let γ_{ik_i}, $(i = 1, 2, \cdots, m)$, be a set of cycles contained in $S(a, \sigma')$. Then $\gamma_{i1} \sim 0$ in $R^n - \bar{G}$; $\gamma_{i1} \sim \gamma_{ik_i} \sim 0$ in $R^n - (M + F(a, \epsilon) - G)$. These homologies determine an $(r + 1)$-cycle γ_i^{r+1}. If γ_i^{r+1} bounds in $R^n - (\bar{G} - G)$, then γ_{i1} bounds in $R^n - (M + F(a, \epsilon))$, or in $S(a, \epsilon) - M$, a contradiction. Consequently γ_i^{r+1} links $\bar{G} - G$, and similar reasoning shows that the m cycles γ_i^{r+1} are independent in $R^n - (\bar{G} - G)$. Consequently $p^{n-r-2}(\bar{G} - G) \geqq m$ and $\beta^{n-r-2}(a, M) \geqq m$.

If $p_a^{n-r-1}(M)$ is infinite, $\beta^{n-r-2}(a, M) = \omega$ or $\aleph_0$, the first case occurring if it is impossible to choose a $\sigma > 0$ so that some γ_{ik_i} lies outside $S(a, \sigma)$, for all i.

Since $p^{n-r-1}(a, M)$ finite or ω is a sufficient condition that M have no $(n - r - 1)$-dimensional condensation at a, we have the

COROLLARY. *If $p^r(a, M)$ is finite or ω, then $\beta^{r-1}(a, M) \geqq p^r(a, M)$.*

[28] See footnote 1, pp. 16 and 25.

14. Unsolved problems. 1. It seems reasonable to suppose that many of the theorems in the theory of order of points of a 1-dimensional set could be extended to the n-dimensional case in terms of the β's. However, in most cases, this seems to be very difficult.

2. Another problem is to give sufficient conditions, in terms of the β's and, probably, local connectedness properties, that a point of an n-dimensional compact metric space have a neighborhood which can be imbedded in R^n.

3. A problem closely related to the preceding is that of extending the characterizations of sections 10, 11, and 12 to the corresponding n-dimensional sets.

4. Finally, under what conditions does the equality $\beta^{i-1}(a, M) = p^i(a, M)$ hold? It seems probable that a partial answer is that it does whenever the latter is finite, but this has yet to be proved.

UNIVERSITY OF MICHIGAN.

ON THE POISSON SUMMABILITY OF FOURIER SERIES

By Norman Levinson

1. Let $f(x)$ be a Lebesgue integrable function of period 2π, and let

$$\phi(x) = f(y + x) + f(y - x) - 2s.$$

It is well known that if

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^\epsilon \left(1 - \frac{x}{\epsilon}\right)^{m-1} \phi(x)\, dx = 0, \tag{1.0}$$

then

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^1 \phi(x)\, dx \int_0^1 (1 - z)^n \cos \frac{xz}{\epsilon}\, dz = 0 \tag{1.1}$$

for $n > m$, where (1.1) is the n-th Riesz mean of the Fourier series for $f(x)$ at $x = y$.

In his conversation class, Hardy carried this relation over to Poisson summability of Fourier series by proving in a very simple manner that

$$\lim_{\epsilon \to 0} \epsilon \int_0^1 \frac{\phi(x)}{x^2} e^{-\frac{\epsilon^2}{x^2}}\, dx = 0$$

implies the Poisson summability of the Fourier series of $f(x)$ at the point $x = y$, and conjectured that

$$\lim_{\epsilon \to 0} \epsilon \int_0^1 \phi(x)\, e^{-\left(\frac{\epsilon}{x}\right)^{1+b}} \frac{dx}{x^2} = 0$$

also implies the P summability for $b > 0$. We shall show this to be the case. We shall also show that there is another exponential kernel $\exp[-(x/\epsilon)^{1+b}]$, similarly related to P summability.

Our theorems are

Theorem 1. *Let $E(m, \alpha)$ represent*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^1 \left(\frac{x}{\epsilon}\right)^\alpha e^{-\left(\frac{x}{\epsilon}\right)^{1+m}} \phi(x)\, dx = 0, \qquad m > -1, \qquad \alpha \geqq 0, \tag{1.2}$$

and $P(m)$ represent

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^1 \frac{\phi(x)}{\left(\frac{x}{\epsilon}\right)^{2(1+m)} + 1}\, dx = 0, \qquad m > -\frac{1}{2}, \tag{1.3}$$

where $\phi(x)$ is defined as above. Then $E(n, \alpha)$ for $n > m$ and $\alpha \geqq 0$, or $E(m, \alpha)$ for $\alpha > 0$ implies $P(m)$, while $P(m)$ implies $E(n, \alpha)$ for $m > n$.

Received October 7, 1935.

THEOREM 2. *If $E^{-1}(m, \alpha)$ represents*

$$\lim_{\epsilon \to 0} \epsilon \int_0^1 \left(\frac{\epsilon}{x}\right)^{\alpha} e^{-\left(\frac{\epsilon}{x}\right)^{1+m}} \phi(x) \frac{dx}{x^2} = 0, \qquad m > -1, \qquad \alpha \geqq 0, \tag{1.4}$$

and $P(m)$ is defined as in Theorem 1, *the conclusion of Theorem* 1 *holds with E replaced by E^{-1}.*

$P(m)$ becomes ordinary Poisson summability for $m = 0$.

Theorem 1 shows that $E(m, \alpha)$ implies $E(n, \beta)$ if $m > n$. That $E(m, \alpha)$ implies $E(m, \beta)$ if $\alpha > \beta \geqq 0$ follows immediately from

$$e^{-x^{1+m}} x^{\beta} = \frac{1+m}{\Gamma\left(\dfrac{\alpha - \beta}{1+m}\right)} \int_0^1 (1 - y^{1+m})^{\frac{\alpha-\beta}{1+m} - 1} y^{\beta} e^{-\left(\frac{x}{y}\right)^{1+m}} \left(\frac{x}{y}\right)^{\alpha} \frac{dy}{y}.$$

Similar results hold for $E^{-1}(m, \alpha)$.

2. Since Wiener's fundamental work[1] on tauberian and related theorems, it is quite natural to use Fourier transform methods on these theorems.[2] We require the following lemmas.

LEMMA 1. *Let*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^1 N_1\left(\frac{x}{\epsilon}\right) \phi(x)\, dx = 0, \tag{2.0}$$

where

$$\int_0^1 |\, \phi(x) \,|\, dx < \infty \tag{2.1}$$

and

$$|\, N_1(x) \,| < A < \infty. \tag{2.2}$$

If $R(x)$ is a function such that

$$\int_0^{\infty} |\, R(x) \,|\, dx < \infty, \tag{2.3}$$

and

$$\int_0^1 |\, R(x) \,| \frac{dx}{x} < \infty, \tag{2.4}$$

and if

$$N_2(x) = \int_0^{\infty} R(y)\, N_1\left(\frac{x}{y}\right) \frac{dy}{y}, \tag{2.5}$$

[1] N. Wiener, *Tauberian theorems*, Annals of Mathematics, vol. 33 (1932), pp. 1–100.

[2] Wiener, loc. cit., has successfully applied these methods to Riesz summability of Fourier series.

then

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_0^1 N_2\left(\frac{x}{\epsilon}\right)\phi(x)\,dx = 0. \tag{2.6}$$

Since the following repeated integral is absolutely convergent, the order of integration can be interchanged, giving

$$\begin{aligned}&\frac{1}{\epsilon}\int_0^\infty R\left(\frac{y}{\epsilon}\right)\frac{dy}{y}\int_0^1 N_1\left(\frac{x}{y}\right)\phi(x)\,dx\\ &\quad= \frac{1}{\epsilon}\int_0^1 \phi(x)\,dx\int_0^\infty R\left(\frac{y}{\epsilon}\right)N_1\left(\frac{x}{y}\right)\frac{dy}{y} = \frac{1}{\epsilon}\int_0^1 \phi(x)\,N_2\left(\frac{x}{\epsilon}\right)dx.\end{aligned} \tag{2.7}$$

Since

$$\left|\frac{1}{y}\int_0^1 \phi(x)\,N_1\left(\frac{x}{y}\right)dx\right| < M < \infty,$$

by (2.0), (2.1), and (2.2), it follows that

$$\begin{aligned}&\left|\frac{1}{\epsilon}\int_0^\infty R\left(\frac{y}{\epsilon}\right)\frac{dy}{y}\int_0^1 N_1\left(\frac{x}{y}\right)\phi(x)\,dx\right|\\ &\quad\leqq \left|\frac{1}{\epsilon}\int_0^\delta R\left(\frac{y}{\epsilon}\right)dy\left\{\frac{1}{y}\int_0^1 N_1\left(\frac{x}{y}\right)\phi(x)\,dx\right\}\right| + M\int_{\delta/\epsilon}^\infty |\,R(y)\,|\,dy.\end{aligned}$$

Since for sufficiently small δ the first term on the right is arbitrarily small independently of ϵ, and since for sufficiently small ϵ, and any fixed δ, the second term is arbitrarily small, we have

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_0^\infty R\left(\frac{y}{\epsilon}\right)\frac{dy}{y}\int_0^1 N_1\left(\frac{x}{y}\right)\phi(x)\,dx = 0,$$

which, combined with (2.7), proves the lemma.

LEMMA 2. *Lemma* 1 *remains valid if* (2.4) *is replaced by* $|\,R(x)\,| < A$, $x < \frac{1}{2}$, *and*

$$\int_0^\infty |\,N_1(x)\,|\,\frac{dx}{x} < \infty.$$

For (2.7) remains true, since

$$\begin{aligned}\int_0^\infty \left|R\left(\frac{y}{\epsilon}\right)\right|\left|N_1\left(\frac{x}{y}\right)\right|\frac{dy}{y} &\leqq A\int_0^{\epsilon/2}\left|N_1\left(\frac{x}{y}\right)\right|\frac{dy}{y} + A\int_{\epsilon/2}^\infty\left|R\left(\frac{y}{\epsilon}\right)\right|\frac{dy}{y}\\ &\leqq A\int_0^\infty |\,N_1(y)\,|\,\frac{dy}{y} + A\int_{1/2}^\infty |\,R(y)\,|\,\frac{dy}{y}.\end{aligned}$$

The remainder of the proof follows as in Lemma 1.

From these lemmas, it is evident that the solution of the integral equation (2.5) for $R(x)$ when $N_1(x)$ and $N_2(x)$ are given is very important. It is to solve this

equation that the Fourier transform, or rather in this case the Mellin transform, is so useful. Of course, the Mellin transform is only a Fourier transform upon which the transformation $y = e^x$ has been performed. We require the following two well-known theorems.

THEOREM A. *If $k(w)$ is analytic in the strip $a \leqq u \leqq b$, where $w = u + iv$, and if*

$$\int_{-\infty}^{\infty} | k(u + iv) |^2 \, dv < A < \infty, \qquad a \leqq u \leqq b,$$

there exists a function

$$F(x) = \operatorname*{l.i.m.}_{A\to\infty} \frac{1}{2\pi i} \int_{-iA+u}^{iA+u} k(w) x^{w-1} \, dw, \qquad a \leqq u \leqq b.$$

Moreover,

$$k(w) = \int_0^{\infty} F(x)\, x^{-w} \, dx, \qquad a < u < b,$$

and

$$\int_0^{\infty} x^{-2u+1} | F(x) |^2 \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} | k(u + iv) |^2 \, dv, \qquad a \leqq u \leqq b.$$

THEOREM B. *If $F(x)$ is a function such that*

$$\int_0^{\infty} | F(x) | \, dx < \infty$$

and if

$$k(iv) = \int_0^{\infty} F(x)\, x^{-iv} \, dx,$$

then

$$F(x) = \lim_{A\to\infty} \frac{1}{2\pi} \int_{-A}^{A} x^{iv-1} k(iv) \left(1 - \frac{|v|}{A}\right) dv$$

almost everywhere.

We can now solve the integral equation (2.5). We have[3]

LEMMA 3. *Let $N_1(x)$ and $N_2(x)$ belong to $L\,(0, \infty)$. Let*

$$k_1(w) = \int_0^{\infty} N_1(x)\, x^{-w} \, dx$$

[3] The condition that $N_1(x)$ and $N_2(x)$ belong to $L(0, \infty)$ in this lemma can be replaced by a variety of other conditions. For example, if for some fixed b, $x^{-b}N_1(x)$ and $x^{-b}N_2(x)$ belong to $L(0, \infty)$ and if $r(w)$ is analytic and belongs to L^2 in the strip $b - \delta \leqq u \leqq b + \delta$, then $x^{-b}R(x)$ belongs to $L(0, \infty)$ and it can readily be shown that $R(x)$ is a solution of the equation (2.5).

and let $k_2(w)$ be similarly defined. Let $r(w) = k_2(w)/k_1(w)$. If $r(w)$ is analytic, and

$$\int_{-\infty}^{\infty} | r(u + iv) |^2 \, dv < M < \infty$$

in the strip, $-\delta \leqq u \leqq \delta$, for some $\delta > 0$, then

$$R(x) = \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{2\pi i} \int_{-iA}^{iA} r(w) \, x^{w-1} \, dw \tag{2.8}$$

is a solution of the integral equation (2.5), *and*

$$\int_0^{\infty} | R(x) | \, dx < \infty . \tag{2.9}$$

That $R(x)$ defined as in (2.8) exists follows from Theorem A. It also follows from this theorem that

$$\int_0^{\infty} | R(x) |^2 \, x^{\pm 2\delta + 1} \, dx < \infty .$$

Using Schwarz's inequality, we get (2.9). To show that $R(x)$ is a solution of the integral equation, we set

$$H(x) = \int_0^{\infty} R(y) \, N_1\!\left(\frac{x}{y}\right) \frac{dy}{y} .$$

Then

$$\int_0^{\infty} | H(x) | \, dx \leqq \int_0^{\infty} | R(y) | \, dy \int_0^{\infty} | N_1(x) | \, dx .$$

Thus $H(x)$ belongs to $L(0, \infty)$. Moreover,

$$\begin{aligned} \int_0^{\infty} H(x) \, x^{-iv} \, dx &= \int_0^{\infty} x^{-iv} \, dx \int_0^{\infty} R(y) N_1\!\left(\frac{x}{y}\right) \frac{dy}{y} \\ &= r(iv) \, k_1(iv) = k_2(iv) . \end{aligned}$$

It follows immediately from Theorem B that $H(x) = N_2(x)$. This proves the lemma.

3. In proving Theorems 1 and 2 we shall need the following transforms, which can readily be computed.

$$\int_0^{\infty} \frac{x^{-w}}{x^{2(1+m)} + 1} \, dx = \frac{\pi}{2(1 + m) \cos \dfrac{\pi}{2} \left(\dfrac{w + m}{1 + m}\right)}, \qquad m > -\frac{1}{2},$$

$$\int_0^{\infty} x^{\alpha - w} \, e^{-x^{1+m}} \, dx = \frac{1}{1 + m} \Gamma\!\left(\frac{1 - w + \alpha}{1 + m}\right), \qquad m > -1, \qquad \alpha \geqq 0,$$

$$\int_0^{\infty} x^{-\alpha - 2 - w} \, e^{-x^{-1-m}} \, dx = \frac{1}{1 + m} \Gamma\!\left(\frac{1 + w + \alpha}{1 + m}\right), \qquad m > -1, \qquad \alpha \geqq 0.$$

We shall also make frequent use of

$$|\Gamma(a+iy)| \sim (2\pi)^{\frac{1}{2}} e^{-\frac{\pi|y|}{2}} |y|^{a-\frac{1}{2}}$$

in what follows.

Proof of Theorem 2. First we prove that $P(m)$ implies $E^{-1}(n, \alpha)$ if $m > n$. Let

$$N_1(x) = \frac{1}{x^{2(1+m)}+1}, \qquad N_2(x) = x^{-\alpha-2} e^{-x^{-1-n}}.$$

Then clearly $N_1(x)$ and $N_2(x)$ are absolutely integrable. Using the terminology of Lemma 3,

$$r(w) = \frac{k_2(w)}{k_1(w)} = \frac{2(1+m)}{\pi(1+n)} \Gamma\left(\frac{1+w+\alpha}{1+n}\right) \cos\frac{\pi}{2}\left(\frac{w+m}{1+m}\right)$$

is analytic and belongs to L^2 uniformly along every ordinate for which $-\delta \leqq u \leqq 1+\delta$ for some $\delta > 0$. Thus the conditions of Lemma 3 are satisfied, and there exists an absolutely integrable $R(x)$ such that (2.5) is satisfied. By Theorem A, setting $u = 1 + \delta$,

$$\int_0^\infty |R(x)|^2 x^{-1-2\delta}\, dx < \infty.$$

Thus by Schwarz's inequality

$$\int_0^1 |R(x)| \frac{dx}{x} < \infty.$$

The conditions of Lemma 1 are now fulfilled, and therefore

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_0^1 \phi(x) N_2\left(\frac{x}{\epsilon}\right) dx = 0.$$

This proves that $P(m)$ implies $E^{-1}(n, \alpha)$ if $m > n$.

We shall now prove that part of the theorem which states that $E^{-1}(n, \alpha)$ implies $P(m)$ for $n > m$. Let

$$N_1(x) = x^{-\alpha-2} e^{-x^{-1-n}}, \qquad N_2(x) = \frac{1}{x^{2(1+m)}+1}.$$

Here

$$r(w) = \frac{k_2(w)}{k_1(w)} = \frac{\pi(1+n)}{2(1+m)} \frac{1}{\Gamma\left(\frac{1+w+\alpha}{1+n}\right) \cos\frac{\pi}{2}\left(\frac{w+m}{1+m}\right)}$$

is analytic and belongs to L^2 uniformly along every ordinate for $-\delta \leqq u \leqq \delta$ for some $\delta > 0$. Thus as before, Lemma 3 is satisfied and an absolutely integrable

$$R(x) = \frac{\pi(1+n)}{2(1+m)} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{x^{w-1}\, dw}{\Gamma\left(\frac{1+\alpha+w}{1+n}\right) \cos\frac{\pi}{2}\left(\frac{w+m}{1+m}\right)} \tag{3.0}$$

exists. If now we displace the path of integration to the right and observe that $w = 1$ is a pole, we have

$$R(x) = \frac{1+n}{1+m} \frac{1}{\Gamma\left(\dfrac{1+m+\alpha}{1+n}\right)} + \frac{\pi(1+n)}{2(1+m)} \frac{1}{2\pi i} \int_{-i\infty+2+2m}^{i\infty+2+2m} \frac{x^{w-1}\, dw}{\Gamma\left(\dfrac{1+\alpha+w}{1+n}\right) \cos \dfrac{\pi}{2} \left(\dfrac{w+m}{1+m}\right)}.$$

But this yields

$$| R(x) | \leqq \frac{1+n}{1+m} \frac{1}{\Gamma\left(\dfrac{1+m+\alpha}{1+n}\right)} + Ax^{1+2m},$$

where A is some positive number. Since $m > -\frac{1}{2}$, $R(x)$ is bounded for finite x and the conditions of Lemma 2 are fulfilled. This proves that $E^{-1}(n, \alpha)$ implies $P(m)$ if $n > m$.

Finally, we want to show that $E^{-1}(m, \alpha)$, $\alpha > 0$, implies $P(m)$. The proof is precisely like the preceding one in every detail, except that $R(x)$ in (3.0) is defined as a limit in the mean. This completes the proof of Theorem 2.

Proof of Theorem 1. First we show that $P(m)$ implies $E(n, \alpha)$ if $m > n$. This proceeds just as in the corresponding proof in Theorem 2 if we observe that

$$r(w) = \frac{2(1+m)}{\pi(1+n)} \Gamma\left(\frac{1-w+\alpha}{1+n}\right) \cos \frac{\pi}{2} \left(\frac{w+m}{1+m}\right)$$

is analytic and belongs to L^2 along every ordinate $-\delta \leqq u \leqq 1 + \delta$ for some $\delta > 0$.

Now let us prove that $E(n, \alpha)$ implies $P(m)$ for $n > m$. If $\alpha = 0$, then

$$r(w) = \frac{\pi(1+n)}{2(1+m)} \frac{1}{\Gamma\left(\dfrac{1-w}{1+n}\right) \cos \dfrac{\pi}{2} \left(\dfrac{w+m}{1+m}\right)},$$

and the proof goes just as in the first part of Theorem 2 by the use of Lemmas 1 and 3. If $\alpha > 0$, then the proof proceeds in the same way as in the second part of Theorem 2. Here

$$R(x) = \frac{\pi(1+n)}{2(1+m)} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{x^{w-1}\, dw}{\Gamma\left(\dfrac{1+\alpha-w}{1+n}\right) \cos \dfrac{\pi}{2} \left(\dfrac{w+m}{1+m}\right)}.$$

We displace the path of integration to the right to the ordinate $u = 1 + \delta$, where $\delta > 0$ is sufficiently small so that the path crosses only the pole at $w = 1$, and in this way, we can show that $R(x)$ is bounded for finite x. In every other detail, the proofs are identical.

Finally, we prove that $E(m, \alpha)$ implies $P(m)$ for $\alpha > 0$. As usual, Lemma 3 leads immediately to an absolutely integrable $R(x)$ defined by

$$R(x) = \frac{\pi}{2} \underset{A\to\infty}{\text{l.i.m.}} \int_{-iA}^{iA} \frac{x^{(w-1)(m+1)}\, dw}{\Gamma\left(1 - w + \frac{\alpha}{1+m}\right) \cos\frac{\pi}{2} w}.$$

Using $\frac{\pi}{\sin \pi w} = \Gamma(w)\, \Gamma(1 - w)$, we have

$$R(x) = (1 + m) \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{2\pi i} \int_{-iA}^{iA} \frac{x^{(m+1)(w-1)}\, \Gamma(1 - w)\, \Gamma(w) \sin \frac{1}{2}\pi w}{\Gamma\left(1 - w + \frac{\alpha}{1+m}\right)}\, dw$$

$$= \frac{1 + m}{\Gamma\left(\frac{\alpha}{1+m}\right)} \frac{1}{2\pi i} \int_{-i\infty-\frac{1}{2}}^{i\infty-\frac{1}{2}} x^{(m+1)(w-1)}\, \Gamma(w) \sin \tfrac{1}{2}\pi w dw \int_0^1 z^{-w}(1 - z)^{\frac{\alpha}{1+m}-1}\, dz$$

$$= \frac{(1 + m)x^{-1-m}}{\Gamma\left(\frac{\alpha}{1+m}\right)} \int_0^1 (1 - z)^{\frac{\alpha}{1+m}-1}\, dz \left[\frac{1}{2\pi i} \int_{-i\infty-\frac{1}{2}}^{i\infty-\frac{1}{2}} \frac{x^{(1+m)w}}{z^w}\, \Gamma(w) \sin \tfrac{1}{2}\pi w dw\right]$$

$$= \frac{(1 + m)x^{-1-m}}{\Gamma\left(\frac{\alpha}{1+m}\right)} \int_0^1 (1 - z)^{\frac{\alpha}{1+m}-1} \sin \frac{z}{x^{1+m}}\, dz.$$

If $\alpha \geqq 1 + m$, it follows from the second law of the mean that $R(x)$ is bounded. Thus for $\alpha \geqq 1 + m$, an application of Lemma 2 proves this part of Theorem 1.

If $\alpha < 1 + m$, we must proceed somewhat differently. We have, setting $z = 1 - yx^{1+m}$,

$$R(x) = \frac{(1 + m)x^{-1-m+\alpha}}{\Gamma\left(\frac{\alpha}{1+m}\right)} \left[\sin\frac{1}{x^{1+m}}\left\{\Gamma\left(\frac{\alpha}{1+m}\right)\cos\frac{\pi\alpha}{2(1+m)} - \int_{x^{-1-m}}^{\infty} \cos y\, y^{\frac{\alpha}{1+m}-1}\, dy\right\}\right.$$

$$\left. - \cos \frac{1}{x^{1+m}} \left\{\Gamma\left(\frac{\alpha}{1+m}\right) \sin \frac{\pi\alpha}{2(1+m)} - \int_{x^{-1-m}}^{\infty} \sin y\, y^{\frac{\alpha}{1+m}-1}\, dy\right\}\right] \tag{3.1}$$

$$= (1 + m)x^{-1-m+\alpha} \sin\left(\frac{1}{x^{1+m}} - \frac{\pi\alpha}{2(1+m)}\right) + O(1).$$

Clearly we cannot use either Lemma 1 or Lemma 2 because of the behavior of $R(x)$ at $x = 0$. However, we can show that the operations performed in these lemmas are valid here. First we show that the inversion of the order of integration of (2.7) is justified in this case. We have

$$\frac{1}{\epsilon} \int_0^\infty R\left(\frac{y}{\epsilon}\right) \frac{dy}{y} \int_0^1 N_1\left(\frac{x}{y}\right) \phi(x)\, dx = \lim_{\delta\to 0} \frac{1}{\epsilon} \int_0^1 \phi(x)\, dx \int_\delta^\infty R\left(\frac{y}{\epsilon}\right) N_1\left(\frac{x}{y}\right) \frac{dy}{y},$$

since $N_1(x) = x^\alpha e^{-x^{1+m}}$ is uniformly bounded and $R(x)$ is absolutely integrable. If we can show that

$$\lim_{\delta \to 0} \int_0^1 \phi(x)\, dx \int_0^\delta R\left(\frac{y}{\epsilon}\right) N_1\left(\frac{x}{y}\right) \frac{dy}{y} = 0, \tag{3.2}$$

we have (2.7) of Lemma 1, and can proceed with the remaining argument of Lemma 1 to complete the proof of this final part of Theorem 1.

Using (3.1), we can prove (3.2), if we can show that

$$I_1 = \lim_{\delta \to 0} \int_0^1 \phi(x)\, dx \int_0^\delta y^{\alpha-1-m} \sin\left(\frac{\epsilon^{1+m}}{y^{1+m}} - \frac{\pi\alpha}{2(1+m)}\right)\left(\frac{x}{y}\right)^\alpha e^{-\left(\frac{x}{y}\right)^{1+m}} \frac{dy}{y} = 0, \tag{3.3}$$

and

$$\lim_{\delta \to 0} \int_0^1 |\, \phi(x)\, |\, dx \int_0^\delta \left(\frac{x}{y}\right)^\alpha e^{-\left(\frac{x}{y}\right)^{1+m}} \frac{dy}{y} = 0. \tag{3.4}$$

But (3.4) can readily be shown by breaking the x interval of integration (0, 1) into $(0, \delta^{\frac{1}{2}})$ and $(\delta^{\frac{1}{2}}, 1)$.

For (3.3), we get, on setting $w = y^{-1-m}$,

$$I_1 = -\frac{1}{1+m} \lim_{\delta \to 0} \int_0^1 x^\alpha \phi(x)\, dx \int_{\delta^{-1-m}}^\infty e^{-wx^{1+m}} \sin\left(\epsilon^{1+m} w - \frac{\pi\alpha}{2(1+m)}\right) dw,$$

and by the second law of the mean

$$|\, I_1 \,| \leqq \frac{2}{(1+m)\epsilon^{1+m}} \lim_{\delta \to 0} \int_0^1 x^\alpha \,|\, \phi(x) \,|\, e^{-\left(\frac{x}{\epsilon\delta}\right)^{1+m}} dx = 0,$$

since $\phi(x)$ is absolutely integrable. This completes the proof of Theorem 1.

Cambridge, England.

CONCERNING THE TRANSITIVE PROPERTIES OF GEODESICS ON A RATIONAL POLYHEDRON

By Ralph H. Fox and Richard B. Kershner

This paper considers geodesics on ordinary polyhedrons[1] in an abstract space. A geodesic on an ordinary polyhedron becomes an ordinary straight line if the sequence of faces belonging to that geodesic is thought of as spread out on a plane. We shall be concerned in what follows only with *rational polyhedrons*, that is, ordinary polyhedrons in which the sum of all angles at any corner is a rational multiple of π. The problem may be considered as an elementary illustration of the 'billard ball' problem considered by Birkhoff in Chapter VI of his Colloquium Publication *Dynamical Systems* and was suggested to us by Wintner. The geometrical condition of rationality defined above is, in the main, the condition on integrability in the sense of Birkhoff or, in the case of a periodic solution, the rationality of the rotation number.

If a direction is moved parallel to itself along any closed curve, which meets no corners, it can only come back to a finite number of positions, for the closed curve can be deformed continuously, without passing over any corners, into another which admits a decomposition into simple loops, each loop consisting of a closed circuit about a corner. Each circuit changes the direction by the sum of the angles about the corner and there is but a finite number of corners.

Now[2] an "Ueberlagerungsfläche" P for the rational polyhedron Π may be defined as follows. Consider an arbitrary but fixed direction[3] on one of the faces of Π and all possible simple curves on Π starting from a fixed point in the interior of that face and not meeting any corners. Let the initial direction be moved parallel to itself along each of these curves. If two such curves have a common end point but different directions there, we consider the two end points to be on different faces. The totality of faces, distinct in this sense, constitutes the *Ueberlagerungsfläche* P. The most important properties of P are the following.

(1) P is a finite polyhedron. For to each face of Π there corresponds a finite number of faces of P.

Received by the Editors of the Annals of Mathematics November 9, 1934, accepted by them, and later transferred to this journal.

[1] Ordinary polyhedrons are meant in the sense of E. Steinitz, *Polyeder und Raumeinteilungen*, Encyklopädie der Mathematischen Wissenschaften, vol. III, Part I_3, p. 15. Geodesics on polyhedrons have been considered, for instance, by P. Stäckel, *Geodätische Linien auf Polyederflächen*, Rend. Circ. Mat. Palermo, vol. 22 (1906), pp. 141–151 and by C. Rodenberg, *Geodätische Linien auf Polyederflächen*, ibid., vol. 23 (1907), pp. 107–125.

[2] For the ideas involved here cf. H. Weyl, *Die Idee der Riemannschen Fläche*, Berlin, 1923.

[3] The particular choice of this direction has, of course, no influence on the construction we are going to make.

(2) Except at the corners of P, there is one and only one direction parallel to any given direction in any point of P.

(3) As a consequence, a geodesic on P, which is not closed, does not meet itself on P.

(4) Except at corners, the structure of P is locally that of the euclidean plane. At corners its local construction is that of a branch point of finite order of a Riemann surface over a euclidean plane.

If a segment of a geodesic is moved continuously parallel to itself in a direction perpendicular to itself, the resulting set of segments of geodesics is called a (geodesical) *strip*. Of course a strip cannot contain a corner of P in its interior; as long as no corners are brought into the interior a strip may be continued, either by *lengthening* the original segment or by *widening* the parallel movement.

The continuation of a geodesic over a corner has not been defined and, indeed, it is not in general[4] possible to define a unique continuation. However, it is always possible to define a right-hand and a left-hand continuation according as the (oriented) geodesic is considered as a limiting position of parallel geodesic segments to the right or to the left of it. If the geodesic meets several corners, we shall always consider as a finite segment of a *singular*[5] geodesic only such a segment that is the limit of a variable non-singular segment.

As far as geodesics are concerned it is P that is fundamental and not Π. Hence we shall really be interested in the behavior of geodesics on P although the theorem which we shall prove will describe indirectly the behavior of geodesics on Π. The following is the theorem.

If P is the Ueberlagerungsfläche of a rational polyhedron and g is a geodesic on P which is not closed, then g is dense on a subset Γ of P which is the closure of a subregion and may be the whole of P. In addition

i. Γ *consists of a finite number of strips;*

ii. Γ *is bounded by a finite number of segments of singular geodesics with corners at both ends;*

iii. *any geodesic of* Γ (*not excluding the singular geodesics in* Γ) *is dense in* Γ. *In particular, if g is dense in P, any parallel geodesic is dense in P.*

If the geodesic is dense on P, then i and ii of the theorem need no proof. Suppose then that g is a non-closed geodesic which is not dense in P. Then there exists at least one strip free of g on P. Let S_h be such a strip, bounded on at least one side by a segment h which is either a segment of g or a limit of segments of g, and which cannot be widened at the side opposite to h. As long as not both end points of h are corners, we can lengthen S_h with the understanding that if a corner appears in its path we must at the same time make S_h narrow enough to get by. This narrowing can only happen a finite number of times, for there is

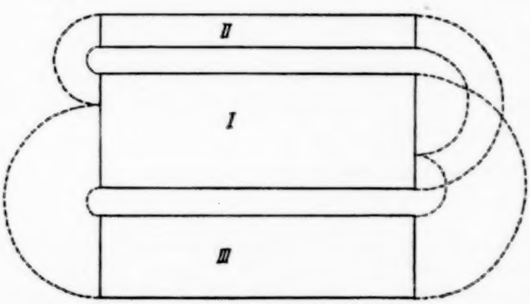
[4] A unique continuation will exist if the sum of the angles about the corner is an integral multiple of π. On the tetrahedron, for example, continuation is unique at all corners, since the sum of the angles about any corner is π.

[5] A singular geodesic is one that meets (at least one) corner. Cf. Stäckel, loc. cit., §12, and Rodenberg, loc. cit., §5.

only a finite number of corners, by (4), and each corner can only be met in a finite number of sheets, by (4), and in each sheet only once, by (2). So the minimum width of S_h is positive and as the area of P is finite, by (1), it follows that the maximum length of h is finite. This cannot be caused by h being closed, for the closure of h implies the closure of g. So the reason must be that both end points of h are corners.

The total number of such strips S_h with maximal h is finite for the total number of segments in the direction of h with corners at both end points is obviously finite. So the width of all these strips has a positive minimum and their length a positive maximum. It is now obvious that the point set Γ covered by g and its closure is the sum of a finite number of strips, and that its boundary consists of a finite number of segments each with corners at both end points.

Let $\bar{g}_1$ be a segment of a geodesic in Γ. Then a geodesic g generated by $\bar{g}_1$ (if g_1 meets a corner in Γ it may actually happen that there are two geodesics generated by $\bar{g}_1$) is a geodesic lying completely in Γ. (It is not excluded that $\bar{g}_1$ be a boundary of Γ, but in this case g_1 consists of the inner continuations of $\bar{g}_1$.) g_1



is not a closed geodesic, since a closed geodesic has a minimum distance from corners and g is by hypothesis dense in Γ. Hence g_1 is dense on a subregion plus its limit points, Γ_1, having properties i and ii. But since every segment of g_1 is a limit of segments of g, we have $\Gamma_1 = \Gamma$. This proves property iii of Γ.

The meaning of the theorem just proved is that, given a direction on a face of Π, P is split up by a finite number of segments of singular geodesics with corners as end points into a (necessarily finite) number of "closures of subregions"

$$\Gamma_1, \Gamma_2, \cdots, \Gamma_\alpha; \qquad \bar{\Gamma}_1, \bar{\Gamma}_2, \cdots, \bar{\Gamma}_\beta; \qquad \alpha \geqq 0, \beta \geqq 0, \alpha + \beta \geqq 1,$$

each "subregion" having properties i and ii and no two "subregions" having any region in common; a geodesic in the given direction in Γ_i lies completely in and is dense on Γ_i; a geodesic in the given direction in $\bar{\Gamma}_i$ lies completely in $\bar{\Gamma}_i$ and is closed (i.e., the $\bar{\Gamma}$'s are strips of closed geodesics).

In general, a direction on a face of Π will not induce an actual "splitting up" of P; what we mean by this is that the number of directions on a face of Π, for

which a geodesic in P in this direction is not everywhere dense on P, is countable. Since P. Stäckel has proved[6] that the number of strips of closed geodesics is countable, it is sufficient to show that the number of segments of singular geodesics with corners at both ends is countable. But this is evident, for the number of such segments, of length less than a given number, is finite.

On the other hand, there do actually exist polyhedrons which "split" into Γ-regions in certain directions. For example, consider a P-surface consisting of the fronts and backs of three rectangles, the width of the first being the sum of the widths of the other two, and joined together as indicated in the diagram on page 149. If the widths of II and III are in an irrational ratio, a geodesic parallel to a side of one of these rectangles will be everywhere dense on the upper sides of these three rectangles.[7] (If the ratio is rational, a geodesic parallel to a side is closed.)

For those rational polygons (two-sided) with which the plane can be paved and those regular polyhedrons whose faces are such polygons (such as the tetrahedron,[8] cube, octahedron, icosahedron, etc., but *not* the dodecahedron) it is certain that non-closed geodesics are dense on the whole Ueberlagerungsfläche. For on them a singular geodesic which meets two corners is necessarily closed, there being but a finite number of sheets and a finite number of corners; hence there cannot exist proper subregions of the type Γ. Whether there exist any proper subregions of the type Γ on the dodecahedron or on an arbitrary rational triangle is an open question.

THE JOHNS HOPKINS UNIVERSITY.

[6] Loc. cit., §10.

[7] The example may be so chosen that it admits a realization in the ordinary three-space. One starts with a rectangle $ABA'B'$ of height 1 and width $1/\sqrt{3}$ and marks points O and O' on AB and $A'B'$, respectively, so that $OB = O'B' = (2 + \sqrt{2})(\sqrt{3} - 1)/2\sqrt{3}$. Subsequently, one marks points P, Q' on AB' and P', Q on $A'B$ so that $\angle AOP = \angle A'O'P' = \pi/6$ and $\angle BOQ = \angle B'O'Q' = \pi/8$. A billiard ball bouncing about the hexagon $POQP'O'Q'$, with initial direction parallel to PQ', will follow the path of a geodesic and will be dense on a certain Γ-region of the Ueberlagerungsfläche of $POQP'O'Q'$. If we fasten a rectangle $RSTU$ to the edge PQ' so that R and S lie between P and Q', a geodesic g, parallel to PQ', will be dense on $POQP'O'Q'STUR$; if parallel to PQ' but below it, g will be closed on $STUR$. This is a concrete example showing that regions of both types, Γ and $\bar{\Gamma}$, may exist on the same polyhedron in the same direction.

[8] The tetrahedron may be treated in a much simpler manner by Kronecker's approximation theorem. For the tetrahedron a non-closed geodesic is not only everywhere dense but uniformly everywhere dense on its Ueberlagerungsfläche. Cf. H. Bohr and R. Courant, *Diophantische Approximationen und Riemannsche ζ-Funktion*, Journal für die reine und angewandte Math., vol. 144 (1914), pp. 258–263.

A CLASS OF BOUNDARY PROBLEMS OF HIGHLY IRREGULAR TYPE

By John I. Vass

1. **Introduction.** In dealing with expansion problems involving ordinary linear differential equations with boundary conditions, Birkhoff,[1] in 1908, distinguished between so-called regular and irregular systems. The distinguishing features of regularity, as defined by him, are intrinsic[2] and not peculiar to the method of treatment. The expansions for arbitrary functions in terms of the solutions of regular systems have been discussed under hypotheses of considerable generality.[3]

The irregularities of differential systems fall into two classifications, mild irregularities and those which are more severe. A system may have one or both types. Expansions involving mildly irregular systems have been treated in rather general cases by Langer[4] and Stone.[5] On the other hand, the expansion problem for highly irregular systems has been successfully studied only in very elementary cases. In so far as is known to the author, this theory has been developed only for differential systems which are highly irregular without being mildly irregular and which have an equation of the type

$$\frac{d^n u}{dx^n} + \lambda^n u = 0, \qquad (n \geqq 3), \tag{1}$$

in the work of Hopkins[6] and L. E. Ward.[7]

If the conditions of regularity are not fulfilled, the series expansions are, in general, non-convergent, but may be summable by suitable means. It was found that a function $f(x)$ satisfies certain very restrictive conditions in order to be expansible as a uniformly convergent series in the solutions of highly

Received August 5, 1935; presented to the American Mathematical Society, April 6, 1934. The author wishes to express his appreciation to Professor R. E. Langer for his many helpful suggestions in the writing of this paper.

[1] G. D. Birkhoff, Trans. Amer. Math. Soc., vol. 9 (1908), pp. 373–395.

[2] D. Jackson, Proceedings Amer. Acad., vol. 51 (1915–16), p. 383, et seq.

[3] J. D. Tamarkin, Mathematische Zeitschrift, vol. 27 (1927), pp. 1–54.

[4] R. E. Langer, Trans. Amer. Math. Soc., vol. 31 (1929), pp. 868–906.

[5] M. H. Stone, Trans. Amer. Math. Soc., vol. 29 (1927), pp. 25–53.

[6] J. W. Hopkins, Trans. Amer. Math. Soc., vol. 20 (1919), pp. 245–259.

[7] L. E. Ward, (a) Trans. Amer. Math. Soc., vol. 29 (1927), pp. 716–731; (b) Annals of Math., vol. 26 (1925), p. 21, et seq.; (c) Trans. Amer. Math. Soc., vol. 34 (1933), pp. 417–434, with the equation $d^3u/dx^3 + [\rho^3 + r(x)]u = 0$, $r(x)$ a convergent power series in x^3.

irregular systems, in which appropriate boundary conditions are associated with equation (1). For instance, in the case of a system involving equation (1) for $n = 3$, Ward[8] showed that, when x is considered complex, the function $f(x)$ is analytic in a specified region of the complex plane; and that, besides satisfying certain other auxiliary conditions, $f(x)$ is of such a nature that when written in the form

$$f(x) \equiv \varphi_1(x^3) + x\varphi_2(x^3) + x^2\varphi_3(x^3),$$

the functions $\varphi_1(x^3)$, $x\varphi_2(x^3)$, $x^2\varphi_3(x^3)$ satisfy certain differential relations dictated by the boundary conditions.

The n linearly independent solutions of equation (1) are

$$e^{\omega^{(2k-1)}\lambda x}, \qquad k = 1, 2, \cdots, n, \qquad \omega = e^{\frac{\pi i}{n}}.$$

The complex numbers $\omega^{(2k-1)}$, $k = 1, 2, \cdots, n$, present in these solutions, are symmetrically located on the unit circle. This symmetry played an essential part in the methods used by both Hopkins and Ward. While for certain purposes such symmetry could be highly desirable, it seems to the author that in these expansion problems it obscures the course of reasoning which must be applied to other problems in which the differential equations have solutions not involving such symmetric numbers.

In the literature no mention is made of the expansion problem for highly irregular systems of the second order. The present paper takes up the expansion theory for such systems of the second order. The reason for this omission in the literature is that with an equation of type (1), $n = 2$, the systems corresponding to those studied by Hopkins and Ward are regular. The highly irregular systems of the second order considered here have an equation with solutions lacking such symmetric numbers as those mentioned above. In the hope that some light may be shed on the general situation for highly irregular boundary problems, the author is here concerned with finding the determining factors for the restrictions on $f(x)$ for systems with an equation differing in type from those previously considered.

The methods used are essentially those of Hopkins and Ward, and the conclusions reached show an appropriate similarity to the results found in their papers. The present paper deduces certain sufficient conditions (see Theorems I and II at the end of this article) for convergence at interior points of the fundamental interval for expansions involving systems which are highly irregular. To avoid complications not pertinent to the argument, the systems considered here are free from mild irregularities.[9]

[8] Reference (7) (a).

[9] A treatment of the expansion problems for analogous systems which are at the same time both mildly and highly irregular may be found in the author's doctor's dissertation on file at the University Library, Madison, Wisconsin.

2. The differential system. Any differential equation of the type

$$u''(x) + m_1\lambda u'(x) + m_2\lambda^2 u(x) = 0,$$

where m_1 and m_2 are such constants, real or complex, that the equation $\gamma^2 + m_1\gamma + m_2 = 0$ has roots which are distinct, non-zero, of equal absolute value, and with arguments differing by a rational fraction of π, can be reduced by a simple change of parameter to the normalized form

$$u''(x) - 2\rho \cos c\, u'(x) + \rho^2 u(x) = 0, \tag{2}$$

with $c = p\pi/q$, $0 < 2p < q$. The functions

$$u_j(x) = e^{\alpha_j \rho x}, \qquad j = 1, 2, \qquad \alpha_1 = e^{ic}, \qquad \alpha_2 = e^{-ic}$$

form a complete set of solutions for this equation. Two distinct differential systems will be formed by adjoining boundary conditions to equation (2). As a first condition, System I will have $u'(0) = 0$ and System II will have $u(0) = 0$, and each system will involve the ordinary linear boundary condition

$$W_2(u) = a_{21}u'(0) + a_{20}u(0) + b_{21}u'(1) + b_{20}u(1) = 0.$$

The fundamental interval can be taken as $(0, 1)$ without loss of generality.

It is at once evident that, with $\nu = 0$ and the constant coefficients a_{ij}, b_{ij} properly chosen, these two systems are included in the following System III, which consists of equation (2) with the boundary conditions

$$W_1(u) = W_{10}(u) + \nu\rho^{-1} W_{11}(u) = 0,$$
$$W_2(u) = W_{20}(u) + W_{21}(u) = 0,$$

where

$$W_{i0}(u) = a_{i1}u'(0) + a_{i0}u(0),$$
$$W_{i1}(u) = b_{i1}u'(1) + b_{i0}u(1), \qquad (i = 1, 2).$$

This system is regular for $\nu \neq 0$.

The two-rowed determinants of the matrix

$$\left\| \begin{matrix} W_{10}(u_1), \ \nu\rho^{-1}W_{11}(u_1), & W_{10}(u_2), \ \nu\rho^{-1}W_{11}(u_2) \\ W_{20}(u_1), \ W_{21}(u_1), & W_{20}(u_2), \ W_{21}(u_2) \end{matrix} \right\|$$

are functions of ρ. If $[A_i]$ is a polynomial in ρ^{-1} of the type

$$[A_i] \equiv A_i + \frac{A_{i1}}{\rho} + \frac{A_{i2}}{\rho^2} + \frac{A_{i3}}{\rho^3} \qquad (i = 0, 1, 2, 3, 4, 5;\ A_i \neq 0),$$

the six independent two-rowed determinants of this matrix may be written as follows:

$$
\begin{aligned}
D_0 &\equiv | W_{i0}(u_1), W_{i0}(u_2) | \equiv \rho[A_0], \\
D_1 &\equiv | (\nu\rho^{-1})^{2-i} W_{i1}(u_1), W_{i0}(u_2) | \equiv \rho^2[A_1] e^{\alpha_1 \rho}, \\
D_2 &\equiv | W_{i0}(u_1), (\nu\rho^{-1})^{2-i} W_{i1}(u_2) | \equiv \rho^2[A_2] e^{\alpha_2 \rho}, \\
D_3 &\equiv | (\nu\rho^{-1})^{2-i} W_{i1}(u_1), (\nu\rho^{-1})^{2-i} W_{i1}(u_2) | \equiv \nu[A_3] e^{(\alpha_1+\alpha_2)\rho}, \\
D_4 &\equiv | W_{i0}(u_1), (\nu\rho^{-1})^{2-i} W_{i1}(u_1) | \equiv \rho^2[A_4] e^{\alpha_1 \rho}, \\
D_5 &\equiv | W_{i0}(u_2), (\nu\rho^{-1})^{2-i} W_{i1}(u_2) | \equiv \rho^2[A_5] e^{\alpha_2 \rho}, \qquad (i = 1, 2).
\end{aligned} \tag{3}
$$

The System III is compatible if ρ is a root of the characteristic equation

$$
\Delta(\rho, \nu) = \begin{vmatrix} W_1(u_1), & W_1(u_2) \\ W_2(u_1), & W_2(u_2) \end{vmatrix} = \sum_{i=0}^{3} D_i = 0. \tag{4}
$$

If any one of the terms in this sum is identically zero, the system is highly irregular. This occurs when $\nu = 0$, for then $D_3 \equiv 0$, and it is this system which is to be studied here. Thus for the two values of the parameter $\nu = 0$ and $\nu = 1$ it is possible to make use of the similarity, such as it is, between the highly irregular and the regular systems. It is for this purpose that the parameter ν is inserted in the system.

3. **The characteristic values and the contours γ_n.** Transcendental equations, such as equation (4), and their solutions have been discussed by Langer[10] and others. From their work it is readily inferred that there are two distributions of characteristic values, one for $\nu = 0$ and one for $\nu = 1$. If $\nu = 1$, these values are spaced at asymptotically regular intervals along the rays I, II, IIIa, and IIIb (see Fig. 1). For the irregular system with $\nu = 0$ the values are spaced along rays I, II, and III, where III is the positive axis of reals.

The characteristic values thus located may be ordered for either set according to numerical magnitude by assigning subscripts so that $|\rho_n| \leqq |\rho_{n+1}|$. If characteristic values have the same numerical magnitude, either actually or asymptotically, they are considered in counter-clockwise order.

It is possible to construct[11] in the ρ-plane an infinite sequence of contours, γ_n, $n = n_1, n_2, n_3, \cdots$, which are concentric circles with center at the origin, the same sequence serving for both ν's. These contours have the following properties.

(i) Every point of any contour γ_n lies at a distance greater than a definite positive constant δ from any characteristic value;

(ii) The contour γ_n has a radius R_n, where $R_n \to \infty$ as $n \to \infty$ over the sequence $n_1, n_2, \cdots$;

[10] R. E. Langer, Trans. Amer. Math. Soc., vol. 31 (1929), pp. 837–844.

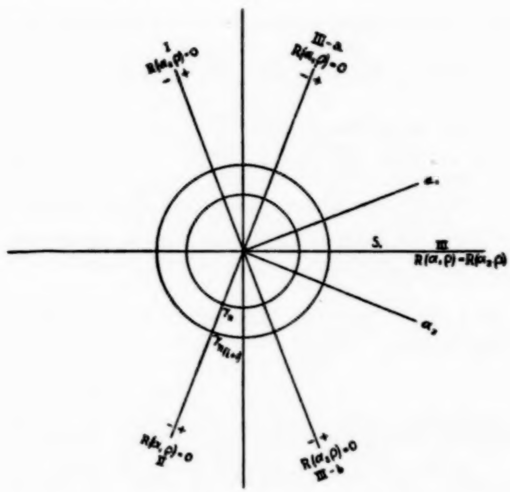
[11] See reference (4), p. 878.

(iii) The contour γ_n contains in its interior just n characteristic values;
(iv) The number r of characteristic values between any two consecutive contours satisfies the inequality $1 \leqq r \leqq 4$.

A further important fact to be noted is that $\Delta(\rho, \nu)$ or the quotient of $\Delta(\rho, \nu)$ divided by any of its terms is uniformly bounded from zero for ρ on any contour γ_n, n sufficiently large. In particular, if $\nu = 0$, for such values of ρ, $\Delta(\rho, 0)$ is expressible in the form

$$\Delta(\rho, 0) \equiv \rho^2 e^{\alpha_1 \rho} \Delta_1(\rho)[A_1], \tag{5}$$

with Δ_1 uniformly bounded from zero.



ρ plane

FIG. 1

The solution of the differential equation (2) which satisfies the first boundary condition is, except possibly for a constant factor,

$$u(x) \equiv W_1(u_2)e^{\alpha_1 \rho x} - W_1(u_1)e^{\alpha_2 \rho x}.$$

This also satisfies the second boundary condition for characteristic values of ρ, thus forming the set of characteristic functions $u^{(i)}(x)$, $i = 1, 2, \cdots$.

4. **The formal expansion of an arbitrary function.** Let $f(x)$ be any arbitrary function and let the function $f_1(x)$ be defined by the relation

$$f_1(x) = f(x) - xf(1) - (1 - x)f(0),$$

so that $f_1(x)$ has the property that it vanishes at both $x = 0$ and $x = 1$. A dis-

cussion will now be made of the formal expansion of $f_1(x)$ in a series of characteristic functions,

$$f_1(x) \sim \sum_{n=1}^{\infty} a_n u^{(n)}(x) . \tag{6}$$

Langer[12] has shown that the sum of the first n terms of such expansions as (6) are associated with the contour integral

$$I_n(x) = \frac{1}{2\pi i} \int_0^1 \int_{\gamma n} M(G) f_1(s) d\rho ds ,$$

where G is the Green's function and the operator $M(G)$ is defined by the relation

$$M(G) = (\alpha_1 + \alpha_2) G_s(x, s, \rho) + \rho G(x, s, \rho) ,$$

the subscript s indicating differentiation with respect to that variable.

The convergence problem will now be taken to be that of showing that the value $f_1(x)$ is the limit of the contour integral $I_n(x)$ as $n \to \infty$ over the sequence $n_1, n_2, n_3, \cdots$. If more than one characteristic value necessarily lies between consecutive contours, the process corresponds to a summation of the series with the terms properly grouped.

5. **The integral $I_n(x)$.** The Green's function for the System III is given by the well-known formula

$$G(x, s, \rho, \nu) = \frac{1}{\Delta(\rho, \nu)} \begin{vmatrix} u_1(x), & u_2(x), & g(x, s, \rho) \\ W_1(u_1), & W_1(u_2), & W_1(g) \\ W_2(u_1), & W_2(u_2), & W_2(g) \end{vmatrix} , \tag{7}$$

where

$$g(x, s, \rho) = \pm \tfrac{1}{2} \frac{u_1(x) u_2(s) - u_2(x) u_1(s)}{u_1'(s) u_2(s) - u_2'(s) u_1(s)} ,$$

the positive or negative sign being taken according as $x >$ or $< s$. The Green's function (7) can be expanded into the form

$$G(x, s, \rho) = g(x, s, \rho) + \frac{1}{\Delta(\rho, \nu)} \left\{ \begin{vmatrix} u_1(x), & u_2(x), & 0 \\ W_{10}(u_1), & W_{10}(u_2), & W_{10}(g) \\ W_2(u_1), & W_2(u_2), & W_2(g) \end{vmatrix} \right.$$

[12] See reference (4), p. 887. The operator $M(G)$ can be derived easily from the integrand of the last integral on the page by the choice of $f_2(x)$ indicated in the footnote there, and integration of the last term by parts.

$$+\frac{\nu}{\rho}\begin{vmatrix} u_1(x), & u_2(x), & 0 \\ W_{11}(u_1), & W_{11}(u_2), & W_{11}(g) \\ W_2(u_1), & W_2(u_2), & W_2(g) \end{vmatrix}\Bigg\}$$

$$= g(x, s, \rho) + h_\nu(x, s, \rho) .$$

Using this latter expression for $G(x, s, \rho)$ in $M(G)$, the integral $I_n(x, \nu)$ becomes

$$(8) \quad I_n(x, \nu) = \frac{1}{2\pi i}\int_0^1 \int_{\gamma_n} M(g) f_1(s) d\rho ds + \frac{1}{2\pi i}\int_0^1 \int_{\gamma_n} M(h_\nu) f_1(s) d\rho ds .$$

From the literature[13] it can be drawn that, if $f_1(x)$ is integrable and of bounded variation on the open interval (0, 1), the first integral on the right of (8) converges to $\frac{1}{2}\{f_1(x^-) + f_1(x^+)\}$ as $n \to \infty$. Likewise from the same source, it is concluded that if $\nu \neq 0$, the second integral on the right of (8) converges uniformly to zero. The proof that this integral converges to zero does not require that $\nu \neq 0$, except for the integration with respect to ρ over γ_{n1}, the arc of γ_n which lies in the sector S_1, where $R(\alpha_1\rho) \geqq 0$ and $R(\alpha\rho) > 0$. (See Fig. 1.) Hence for $\nu = 0$ it is the integral

$$I_{n1}(x) = \frac{1}{2\pi i}\int_0^1 \int_{\gamma_{n1}} M(h_0) f_1(s) d\rho ds ,$$

which requires special treatment and any conditions under which this integral converges uniformly to zero give the results sought. The convergence of this integral is the prime contribution of this paper.

6. **The form of $I_{n1}(x)$.** Since x is present only as a parameter in

$$H_0(x, s, \rho) = \frac{1}{\Delta(\rho)}\begin{vmatrix} u_1(x), & u_2(x), & 0 \\ W_{10}(u_1), & W_{10}(u_2), & W_{10}(g) \\ W_2(u_1), & W_2(u_2), & W_2(g) \end{vmatrix} ,$$

it follows that

$$M(h_0) = \frac{1}{\Delta(\rho)}\begin{vmatrix} u_1(x), & u_2(x), & 0 \\ W_{10}(u_1), & W_{10}(u_2), & W_{10}(M(g)) \\ W_2(u_1), & W_2(u_2), & W_2(M(g)) \end{vmatrix} .$$

Let the function $g(x, s, \rho)$ be written in the form

$$g(x, s, \rho) = \pm \tfrac{1}{2} \sum_{i=1}^{2} u_i(x)\, v_i(s) ,$$

[13] Cf. reference (4).

the positive sign being used if $x > s$ and the negative sign if $x < s$. From differentiation with respect to s of the functions

$$v_1(s) = \frac{e^{-\alpha_1 \rho s}}{\rho(\alpha_1 - \alpha_2)}, \qquad v_2(s) = \frac{-e^{-\alpha_2 \rho s}}{\rho(\alpha_1 - \alpha_2)},$$

the formula for $M(g)$ is found to be

$$M(g) = \mp \tfrac{1}{2}\rho \sum_{i=1}^{2} \alpha_i^2 u_i(x) v_i(s),$$

where the sign is negative or positive according as $x >$ or $< s$. In similar fashion,

$$W_{10}[M(g)] = \tfrac{1}{2}\rho \sum_{i=1}^{2} \alpha_i^2 v_i(s) W_{10}(u_i),$$

$$W_2[M(g)] = \tfrac{1}{2}\rho \sum_{i=1}^{2} \alpha_i^2 v_i(s) [W_{20}(u_i) - W_{21}(u_i)].$$

Thus

$$M(h_0) = \frac{1}{\Delta(\rho)} \begin{vmatrix} u_1(x), & u_2(x), & 0 \\ W_{10}(u_1), & W_{10}(u_2), & \tfrac{1}{2}\rho \sum_{i=1}^{2} \alpha_i^2 v_i(s) W_{10}(u_i) \\ \sum_{i=0}^{1} W_{2i}(u_1), & \sum_{i=0}^{1} W_{2i}(u_2), & \tfrac{1}{2}\rho \sum_{i=1}^{2} \alpha_i^2 v_i(s) [W_{20}(u_i) - W_{21}(u_i)] \end{vmatrix},$$

and by linear combination of columns

$$M(h_0) = \frac{1}{\Delta(\rho)} \begin{vmatrix} u_1(x), & u_2(x), & R(x, s, \rho) \\ W_{10}(u_1), & W_{10}(u_2), & \sum_{i=1}^{2} \rho\alpha_i^2 v_i(s) W_{10}(u_i) \\ W_2(u_1), & W_2(u_2), & \sum_{i=1}^{2} \rho\alpha_i^2 v_i(s) W_{20}(u_i) \end{vmatrix},$$

where

$$R(x, s, \rho) = \tfrac{1}{2}\rho \sum_{i=1}^{2} \alpha_i^2 u_i(x) v_i(s).$$

This evaluation of $M(h_0)$ may be used in the integral $I_{n1}(x)$. The determinant displayed in $M(h_0)$ is now expanded by means of the minors of the elements in the last column. The integrals due to the term $R(x, s, \rho)$ will not be discussed further, since they may be merged with the first term on the right of (8). If the symbols y_{ij} are defined by

$$y_{ij}(x, s) = \alpha_j^2 u_i(x) v_i(s), \qquad (i, j = 1, 2),$$

the contribution I'_{n1} due to $M(h_0)$ becomes

$$I'_{n1}(x) = \frac{-1}{2\pi i}\int_0^1\int_{\gamma_{n1}}\frac{\rho}{\Delta(\rho)}\{(D_0 + D_2)y_{11} + D_3y_{12} - D_4y_{21} + (D_0 + D_1)y_{22}\}\,f_1(s)d\rho ds\,,$$

where the D_j are those of (3) (with $\nu = 0$).

7. The convergence of the integral $I'_{n1}(x)$. The integral $I'_{n1}(x)$ will now be expanded by substituting the definitions of $u_i(x)$, $v_i(s)$ in the functions $y_{ij}(x, s)$. If $l_i = \dfrac{\alpha_i}{\alpha_1 - \alpha_2}$, $i = 1, 2$, this integral can be expressed as the sum of six integrals,

$$\text{(9)}\qquad\begin{aligned}
&\text{(a)}\quad -\frac{\alpha_1 l_1}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta^{-1}(\rho)e^{\alpha_1\rho(x-s)}\,\rho[A_0]f_1(s)\,d\rho ds\,,\\
&\text{(b)}\quad -\frac{\alpha_1 l_1}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta^{-1}(\rho)e^{\alpha_1\rho(x-s)+\alpha_2\rho}\,\rho^2[A_2]f_1(s)\,d\rho ds\,,\\
&\text{(c)}\quad \frac{\alpha_2 l_2}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta^{-1}(\rho)e^{\alpha_1\rho x+\alpha_2\rho(1-s)}\,\rho^2[A_3]f_1(s)\,d\rho ds\,,\\
&\text{(d)}\quad \frac{\alpha_1 l_1}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta^{-1}(\rho)e^{\alpha_2\rho x+\alpha_1\rho(1-s)}\,\rho^2[A_4]f_1(s)\,d\rho ds\,,\\
&\text{(e)}\quad \frac{\alpha_2 l_2}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta^{-1}(\rho)e^{\alpha_2\rho(x-s)}\,\rho[A_0]f_1(s)\,d\rho ds\,,\\
&\text{(f)}\quad \frac{\alpha_2 l_2}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta^{-1}(\rho)e^{\alpha_2\rho(x-s)+\alpha_1\rho}\,\rho^2[A_1]f_1(s)\,d\rho ds\,.
\end{aligned}$$

The sector S_1 is now to be divided into the sub-sectors S_{1j}, $j = 1, 2$, equal in size and with the positive axis of reals as separating boundary. For the sector S_{11}, which is situated in the fourth quadrant, $0 < R(\alpha_2\rho) \leqq R(\alpha_1\rho)$, while for the sector S_{12} in the first quadrant, $0 \leqq R(\alpha_1\rho) < R(\alpha_2\rho)$. The convergence proof will be carried out for the sector S_{11} only, since the argument is the same for S_{12} with the rôles played by α_1 and α_2 interchanged. Hereafter, the symbol γ_{n1} will be understood to refer to the arc of γ_{n1} in S_{11}.

The integrals (a) and (e) in (9) will be considered first. If the form (5) is used for $\Delta(\rho)$ in these integrals, they become

$$\begin{aligned}
&\text{(a)}\quad \frac{-\alpha_1 l_1}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta_1^{-1}(\rho)e^{\alpha_1\rho(x-s-1)}\left[\frac{A_0}{A_1}\right]f_1(s)\,\frac{d\rho}{\rho}\,ds\,,\\
&\text{(e)}\quad \frac{\alpha_2 l_2}{2\pi i}\int_0^1\int_{\gamma_{n1}}\Delta_1^{-1}(\rho)e^{\alpha_2\rho(x-s-1)-(\alpha_1\rho-\alpha_2\rho)}\left[\frac{A_0}{A_1}\right]f_1(s)\,\frac{d\rho}{\rho}\,ds\,,
\end{aligned}$$

respectively. In this form these two integrals are readily seen to converge uniformly to zero by reason of the following lemma.

LEMMA I.[14] *Let ξ be a real variable on the interval $(0, \xi_1)$, ρ a complex variable ranging over the circular contours γ_n, and γ_n' any arc of γ_n lying entirely in the half-plane $R(c_1\rho) \leqq 0$ (c_1 a constant $\neq 0$). Then, if $E(\rho)$ is uniformly bounded for $|\rho|$ sufficiently large and $\psi(\xi)$ is any function which is integrable on the interval $(0, \xi_1)$ and $0 \leqq \xi' \leqq \xi'' \leqq \xi_1$, the integral*

$$\int_{\xi'}^{\xi''} \int_{\gamma_n'} E(\rho) e^{c_1 \rho \xi} \psi(\xi) \frac{d\rho}{\rho} \, d\xi \to 0$$

uniformly as $n \to \infty$.

The irregularity of the problem is such that the convergence for the remaining four integrals of (9) cannot be shown by methods like those employed for the integrals (a) and (e). The individual integrals apparently do not converge, and hence combinations of them are to be considered. Let it be assumed that $f_1(x)$ is differentiable. Then, since $f_1(s)$ vanishes at both $s = 0$ and $s = 1$, integrations by parts with respect to s leave the integrals in the forms

$$(10) \qquad \begin{aligned} &\text{(a)} \quad \frac{l_1}{2\pi i} \int_0^1 \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_1 \rho (x-s) + \alpha_2 \rho} \rho [A_2] f_1'(s) \, d\rho ds, \\ &\text{(b)} \quad \frac{-l_2}{2\pi i} \int_0^1 \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_1 \rho x + \alpha_2 \rho (1-s)} \rho [A_5] f_1'(s) \, d\rho ds, \\ &\text{(c)} \quad \frac{-l_1}{2\pi i} \int_0^1 \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_2 \rho x + \alpha_1 \rho (1-s)} \rho [A_4] f_1'(s) \, d\rho ds, \\ &\text{(d)} \quad \frac{-l_2}{2\pi i} \int_0^1 \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_2 \rho (x-s) + \alpha_1 \rho} \rho [A_1] f_1'(s) \, d\rho ds. \end{aligned}$$

For these the parts arising from integrating from x to 1 can be shown to converge to zero by the methods used for the integrals (a) and (e).

8. **The integrals** (10) **over the interval** $(0, x)$. It remains to be shown that the portions of the integrals (10) taken over the interval $(0, x)$ are uniformly convergent to zero as $n \to \infty$ when $f_1(s)$ is suitably restricted. To show this convergence, a change in the path of integration with respect to s is made. As a first step, in the integrals involving $e^{-\alpha_1 \rho s}$ the change of variable $s = \alpha_2 t/\alpha_1$ is made, while in the integrals involving $e^{-\alpha_2 \rho s}$, s is merely replaced by t. This has the effect of making the exponents in these two exponentials the same, so that the integrals become

$$(11) \qquad \begin{aligned} &\frac{l_2}{2\pi i} \int_0^{\alpha_1 x/\alpha_2} \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_1 \rho x + \alpha_2 \rho (1-t)} \rho [A_2] f_1'\left(\frac{\alpha_2}{\alpha_1} t\right) d\rho dt, \\ &\frac{-l_2}{2\pi i} \int_0^{x} \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_1 \rho x + \alpha_2 \rho (1-t)} \rho [A_5] f_1'(t) \, d\rho dt, \\ &\frac{-l_2}{2\pi i} \int_0^{\alpha_1 x/\alpha_2} \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_2 \rho (x-t) + \alpha_1 \rho} \rho [A_4] f_1'\left(\frac{\alpha_2}{\alpha_1} t\right) d\rho dt, \\ &\frac{-l_2}{2\pi i} \int^{x} \int_{\gamma_{n1}} \Delta^{-1}(\rho) e^{\alpha_2 \rho (x-t) + \alpha_1 \rho} \rho [A_1] f_1'(t) \, d\rho dt. \end{aligned}$$

[14] Tamarkin, loc. cit., p. 43.

Consider the circle of radius x with center at $x = 0$, which will be referred to as the "x-circle". The path of integration is now changed from the straight lines 0 to x and 0 to $\alpha_1 x/\alpha_2$ to the straight line 0 to $\alpha_1 x$, together with the arcs of the "x-circle" from $\alpha_1 x$ to x and $\alpha_1 x$ to $\alpha_1 x/\alpha_2$. (See Fig. 2.) To render this and subsequent steps valid, the function $f_1(x)$, with x a complex variable, is to be assumed analytic within and on the "x-circle". Furthermore, and this is an essential restriction, $f_1(t)$ is to be such a function that the integrals (11) over the path 0 to $\alpha_1 x$ exactly cancel.

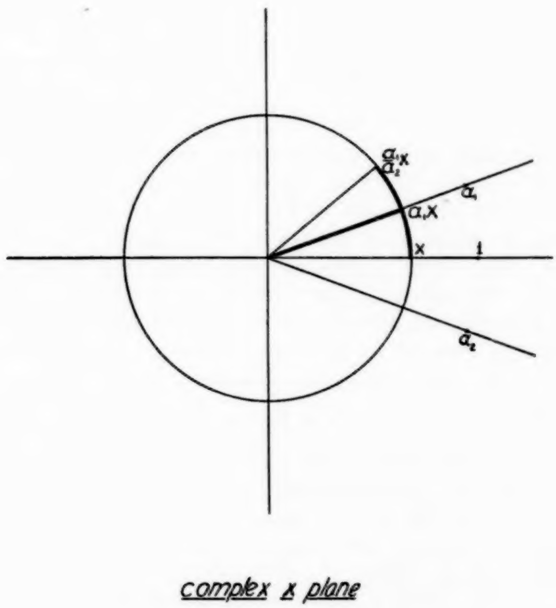


FIG. 2

9. **Convergence along the arcs $\alpha_1 x$ to x and $\alpha_1 x$ to $\alpha_1 x/\alpha_2$.** The two exponentials in the integrals (11) are related by the equality

$$e^{\alpha_1\rho x+\alpha_2\rho(1-t)} \equiv e^{(\alpha_1\rho-\alpha_2\rho)(x-1)}\, e^{\alpha_2\rho(x-t)+\alpha_1\rho}. \tag{12}$$

Since $R(\alpha_1\rho) \geqq R(\alpha_2\rho)$ on the arc γ_{n1} and x is an interior point of the interval (0, 1), the first exponential factor on the right is bounded for large values of $|\rho|$. By means of (12) and the relation

$$\frac{e^{\alpha_2\rho(x-t)+\alpha_1\rho}}{\Delta(\rho)} = \frac{e^{\alpha_2\rho(x-t)}}{\rho^2[A_1]\Delta_1(\rho)}$$

the portion of the integrals (11) over the arcs $\alpha_1 x$ to $\alpha_1 x/\alpha_2$ and $\alpha_1 x$ to x can be written in the following form:

$$
(13)\quad
\begin{aligned}
&\text{(a)}\quad \frac{l_2}{2\pi i}\int_{\alpha_1 x}^{\alpha_1 x/\alpha_2}\int_{\gamma_{n1}} E(x,\rho)\, e^{\alpha_2\rho(x-t)}\left[\frac{A_2}{A_1}\right] f_1'\left(\frac{\alpha_2}{\alpha_1}t\right)\frac{d\rho}{\rho}\, dt\,,\\
&\text{(b)}\quad \frac{-\,l_2}{2\pi i}\int_{\alpha_1 x}^{x}\int_{\gamma_{n1}} E(x,\rho)\, e^{\alpha_1\rho(x-t)}\left[\frac{A_5}{A_1}\right] f_1'(t)\frac{d\rho}{\rho}\, dt\,,\\
&\text{(c)}\quad \frac{-\,l_2}{2\pi i}\int_{\alpha_1 x}^{\alpha_1 x/\alpha_2}\int_{\gamma_{n1}} E(\rho)\, e^{\alpha_2\rho(x-t)}\left[\frac{A_4}{A_1}\right] f_1'\left(\frac{\alpha_2}{\alpha_1}t\right)\frac{d\rho}{\rho}\, dt\,,\\
&\text{(d)}\quad \frac{-\,l_2}{2\pi i}\int_{\alpha_1 x}^{x}\int_{\gamma_{n1}} E(\rho)\, e^{\alpha_2\rho(x-t)} f_1'(t)\frac{d\rho}{\rho}\, dt\,,
\end{aligned}
$$

where $E(x, \rho) = \Delta_1^{-1}(\rho)\, e^{(\alpha_1\rho-\alpha_2\rho)(x-1)}$ and $E(\rho) = \Delta_1^{-1}(\rho)$. For these integrals to converge uniformly to zero, it is necessary that along the arc of the "x-circle" x to $\alpha_1 x/\alpha_2$, $R[-\alpha_2\rho(t - x)] \leqq 0$. To show that this is true, let ω be the argument of t so that $(t - x) = x(e^{i\omega} - 1)$. If $R[-\alpha_2\rho(t - x)]$ is to be less than or equal to zero, the argument $\eta = \arg\,(e^{i\omega} - 1)$, considered as the angle of rotation applied to the vector ρ in the complex plane, must be such that the vector ρ remains in the half-plane where $R(\alpha_2\rho) \geqq 0$. Since the vector ρ is in the sector S_{11} bounded by the axis of reals and the ray $R(\alpha_2\rho) = 0$, it must not be rotated positively through an angle greater than $\frac{1}{2}\pi + \arg\alpha_1$. (See Figs. 1 and 2.) Hence η is restricted by the inequality,

$$
(14)\qquad 0 \leqq \eta \leqq \tfrac{1}{2}\pi + \arg\alpha_1\,.
$$

The vector $e^{i\omega} - 1$ is a chord of the unit circle subtending the angle ω at the origin. From the triangle thus formed, it is evident that $\eta = \frac{1}{2}\pi + \frac{1}{2}\omega$. As ω varies from 0 to $\arg\,(\alpha_1/\alpha_2) = 2\arg\alpha_1$, η varies from $\frac{1}{2}\pi$ to $\frac{1}{2}\pi + \arg\alpha_1$. Therefore the conclusion is drawn that the condition (14) is satisfied when t ranges over the arc x to $\alpha_1 x/\alpha_2$.

The fact that the integrals (13) converge uniformly to zero depends on the following lemma. Here E will denote a bounded function of ρ or of x and ρ, while t' and t'' will represent two points on the "x-circle" such that $0 \leqq \arg t' \leqq \arg t'' \leqq 2\arg\alpha_1$.

LEMMA II. *If $\psi(t)$ is integrable on the arc of the "x-circle" for which $0 \leqq \arg t \leqq 2\arg\alpha_1$, the integral*

$$
(15)\qquad I = \int_{t'}^{t''}\int_{\gamma_{n1}} E\, e^{-\alpha_2\rho(t-x)}\,\psi(t)\,\frac{d\rho}{\rho}\, dt \to 0
$$

uniformly as $n \to \infty$.

Proof. With ρ on the arc γ_{n1}, $-\alpha_2\rho$ lies on the arc of the circle γ_n in the second quadrant bounded by the positive axis of imaginaries and the ray through $-\alpha_2$. If φ is defined by the relation $\lambda = -\alpha_2\rho = iR_n e^{i\varphi}$, then φ must satisfy the inequality $0 < \varphi \leqq \frac{1}{2}\pi - \arg\alpha_1$ in the λ-plane.

This change of variable, together with the fact that $t = xe^{i\omega}$, $0 \leqq \omega \leqq 2 \arg \alpha_1$, makes it possible to write the integral (15) in the form

$$I = -\int_{\arg t'}^{\arg t''} \int_0^{\frac{\pi}{2} - \arg \alpha_1} E\, e^{iR_n x(\cos \varphi + i \sin \varphi)\, [(\cos \omega - 1) + i \sin \omega]}\, \psi_1(\omega) x\, e^{i\omega}\, d\varphi d\omega\,.$$

The real part of the product

$$i\,(\cos \varphi + i \sin \varphi)\,[(\cos \omega - 1) + i \sin \omega]$$

is

$$-\,\{\sin \varphi\,(\cos \omega - 1) + \sin \omega \cos \varphi\}\,.$$

By means of the trigonometric relations

$$\cos \omega - 1 = -2 \sin^2 \tfrac{1}{2}\omega\,, \qquad \sin \omega = 2 \sin \tfrac{1}{2}\omega \cos \tfrac{1}{2}\omega$$

this expression reduces to

$$-2 \sin \tfrac{1}{2}\omega \cos (\tfrac{1}{2}\omega + \varphi)\,. \tag{16}$$

The relations

(i) $$\sin \tfrac{1}{2}\omega \geqq \tfrac{1}{4}\omega\,, \qquad 0 \leqq \tfrac{1}{2}\omega \leqq \tfrac{1}{2}\pi\,,$$

(ii) $$\begin{aligned} \cos (\tfrac{1}{2}\omega + \varphi) &= \sin [\tfrac{1}{2}\pi - (\tfrac{1}{2}\omega + \varphi)] \\ &\geqq \sin [\tfrac{1}{2}\pi - (\arg \alpha_1 + \varphi)] \\ &\geqq \tfrac{1}{2}(\tfrac{1}{2}\pi - \arg \alpha_1 - \varphi)\,, \qquad \tfrac{1}{2}\pi \geqq \tfrac{1}{2}\pi - \arg \alpha_1 - \varphi \geqq 0 \end{aligned}$$

are then used to show that the quantity (16) is less than or equal to the value

$$-\,\tfrac{1}{4}\omega\,(\tfrac{1}{2}\pi - \arg \alpha_1 - \varphi)\,.$$

It follows that

$$\begin{aligned} |\,I\,| &\leqq M_1 \int_{\arg t'}^{\arg t''} \int_0^{\frac{1}{2}\pi - \arg \alpha_1} e^{-\frac{1}{4} R_n x\,\omega(\frac{1}{2}\pi - \arg \alpha_1 - \varphi)}\, |\,\psi_1(\omega)\,|\, d\varphi d\omega \\ &\leqq M_2 \int_{\arg t'}^{\arg t''} \frac{4}{R_n x \omega}\, \{1 - e^{-\frac{1}{4} R_n x\,\omega(\frac{1}{2}\pi - \arg \alpha_1)}\}\, |\,\psi_1(\omega)\,|\, d\omega\,. \end{aligned}$$

Therefore

$$|\,I\,| \leqq M_3 \left\{ \frac{1}{R_n \delta_n} \int_0^{2 \arg \alpha_1} |\,\psi_1(\omega)\,|\, d\omega + \int_0^{\delta_n} |\,\psi_1(\omega)\,|\, d\omega \right\} \to 0\,,$$

if δ_n tends to zero, so that $R_n \delta_n \to \infty$ as $n \to \infty$.

10. **Analysis of the restrictions on $f_1(t)$.** Finally, the integrals of (11) exactly cancel along the straight line path 0 to $\alpha_1 x$ if the function $f_1(t)$ is properly restricted. This condition on $f_1(t)$ is of importance, since it determines a class of functions for which the expansion (6) is possible.

It will be observed that the integrals under consideration will cancel over the path 0 to $\alpha_1 x$ when the equality

$$\{e^{\alpha_1\rho x+\alpha_2\rho}\rho^2[A_2] - e^{\alpha_2\rho x+\alpha_1\rho}\rho^2[A_4]\}f_1'\left(\frac{\alpha_2}{\alpha_1}t\right)$$
$$= \{e^{\alpha_1\rho x+\alpha_2\rho}\rho^2[A_5] + e^{\alpha_2\rho x+\alpha_1\rho}\rho^2[A_1]\}f_1'(t)$$

is satisfied. In other notation, this equation reads

$$[u_1(x)D_2 - u_2(x)D_4]f_1'\left(\frac{\alpha_2}{\alpha_1}t\right) = [u_1(x)D_5 + u_2(x)D_1]f_1'(t),$$

or further,

$$W_{10}(u_1)\{u_1(x)W_{21}(u_2) - u_2(x)W_{21}(u_1)\}f_1'\left(\frac{\alpha_2}{\alpha_1}t\right)$$
$$= W_{10}(u_2)\{u_1(x)W_{21}(u_2) - u_2(x)W_{21}(u_1)\}f_1'(t).$$

This equation will be an identity only when

$$W_{10}(u_1)f_1'\left(\frac{\alpha_2}{\alpha_1}t\right) \equiv W_{10}(u_2)f_1'(t)\,. \tag{17}$$

Since $f_1(t)$ is analytic within the unit circle, and $f_1(0) = 0$, this function can be expressed as a power series,

$$f_1(t) \equiv c_1t + c_2t^2 + c_3t^3 + \cdots.$$

For the System I, the condition (17) after integration reduces to

$$f_1\left(\frac{\alpha_2}{\alpha_1}t\right) = \left(\frac{\alpha_2}{\alpha_1}\right)^2 f_1(t), \tag{18}$$

and upon comparison of the power series for the two members of this equality it is seen that equation (18) is an identity if $f_1(t)$ is a function of the type

$$f_1(t) \equiv t^2\sum_{r=0}^{\infty} c_{rk+2}t^{rk}, \tag{19}$$

where the integer k is determined by the relation $(\alpha_2/\alpha_1)^k = 1$. That such an integer k exists follows from the fact that α_1 and α_2 are both commensurable with π.

On the other hand, for the System II, the condition (17) becomes

$$f_1\left(\frac{\alpha_2}{\alpha_1}t\right) \equiv \frac{\alpha_2}{\alpha_1}f_1(t),$$

which is satisfied if

$$f_1(t) \equiv t\sum_{r=0}^{\infty} c_{rk+1}t^{rk}. \tag{20}$$

From the above discussion it is apparent that the actual nature of the functions in (19) and (20) depends upon the value of α_1, that is, upon the argument c in the differential equation. If we return to the definition of $f_1(x)$, it is also clear that a class of functions $f(x)$ which are expansible in a uniformly convergent series of characteristic functions has been determined for each of the Systems I and II, namely: for System I,

$$f(x) = C_1 + xC_2 + x^2\Phi_1(x^{rk}),$$

and for System II,

$$f(x) = C_1 + x\Phi_2(x^{rk}).$$

11. **Sufficient conditions for uniform convergence.** The conditions just established are sufficient to insure uniform convergence of an expansion of a function $f(x)$ in a series of characteristic values for the systems discussed. These conditions are summarized by the following theorems.

THEOREM I. *For the System* I, *sufficient conditions for the expansibility of a function* $f(x)$ *in a uniformly convergent series of characteristic functions on the open interval* (0, 1) *are*

(i) *that the function* $f(x)$ *be integrable and of bounded variation on* $0 < x < 1$,

(ii) *that the function* $f(x)$ *be analytic within the unit circle in the complex plane,*

(iii) *that the function* $f(x)$ *be of the structure*

$$C_1 + xC_2 + x^2\Phi_1(x^{rk}),$$

where k *is determined as the smallest integer for which* $(\alpha_2/\alpha_1)^k = 1$.

THEOREM II. *For the System* II, *sufficient conditions for the expansibility of a function* $f(x)$ *in a uniformly convergent series of characteristic functions on the open interval* (0, 1) *are*

(i) *that the function* $f(x)$ *be integrable and of bounded variation on* $0 < x < 1$,

(ii) *that the function* $f(x)$ *be analytic within the unit circle in the complex plane,*

(iii) *that the function* $f(x)$ *be of the structure*

$$C_1 + x\Phi_2(x^{rk}),$$

where k *is determined as the smallest integer for which* $(\alpha_2/\alpha_1)^k = 1$.

UNIVERSITY OF WISCONSIN AT MILWAUKEE.

A PARTICULAR SEQUENCE OF STEP FUNCTIONS

By Nelson Dunford

1. **Introduction.** If a sequence of real functions $f_n(t)$ summable on (0, 1) converges in measure to $f(t)$ and if

$$\lim_{n=\infty} \int_\delta f_n(t)\,dt \text{ exists} \tag{1}$$

for every measurable subset δ of (0, 1), then $f(t)$ is summable and

$$\lim_{n=\infty} \int_\delta f_n(t)\,dt = \int_\delta f(t)\,dt$$

uniformly with respect[1] to δ. But (1) may hold for every interval δ in (0, 1) and the conclusion in the weaker form

$$\lim_{n=\infty} \int_0^x f_n(t)\,dt = \int_0^x f(t)\,dt \quad \text{almost everywhere}$$

may not be true even if it is assumed that $f(t)$ is summable and that

(2) $f_n(t) = f(t)$, (except on a set whose measure approaches zero with $1/n$).

The present note is concerned with the behavior of $\int_0^x f_n(t)\,dt$ under the assumption (2) and we assume without loss of generality that $f(t) = 0$. The result is that *there exists a sequence of positive step functions $f_n(t)$ satisfying* (2) *with $f(t) = 0$ such that for every summable function $g(t)$ except for those in a certain set of the first category the sequence $\int_0^x f_n(t)g(t)\,dt$ is everywhere dense in the space of measurable functions.* This is embodied in Theorem 2. The principle (which is Theorem 1) underlying the construction of the sequence $f_n(t)$ is a generalization of an abstraction of an argument used by J. Marcinkiewicz[2] to show the existence of a continuous function $\varphi(t)$ which depends only upon a given sequence of positive numbers $h_n \to 0$ such that an arbitrary measurable func-

Received January 19, 1936.

[1] This can be proved by combining results found in the following references. Saks, *Addition to the note on some functionals*, Trans. Amer. Math. Soc., vol. 35 (1933), p. 969; Jeffery, *The integrability of a sequence of functions*, Trans. Amer. Math. Soc., vol. 33 (1931), p. 435, B; and Dunford, *Integration in general analysis*, Trans. Amer. Math. Soc., vol. 37 (1935), p. 447, Theorem 9.

[2] *Sur les nombres dérivés*, Fund. Math., vol. 24 (1935), pp. 305–308.

tion may be approached almost everywhere by an appropriately chosen subsequence of $[\varphi(t + h_n) - \varphi(t)]/h_n$.

2. The fundamental principle underlying the construction.

THEOREM 1. *Let X and Y be metric spaces and $\{y_m\}$ dense in Y. Let $T_n x$ be a sequence of continuous functions with domain X and range in Y satisfying the following condition. For each m there is an unbounded sequence of integers n_i such that the set of x which satisfy the equation $\lim_{i\to\infty} T_{n_i} x = y_m$ is dense in X. Then for every x in X, except for a set of the first category, the sequence $T_n x$ is dense in Y.*

Let X_1 be the set for which $T_n x$ is dense in Y. Then $x \in X_1$ if and only if for every pair of integers m, n there is a $q \geqq n$ such that $(T_q x, y_m) < 1/n$. Thus $X - X_1 = \sum_{m,n} X_{mn}$, where X_{mn} is the set of x for which $(T_q x, y_m) \geqq 1/n$, $q \geqq n$. Since $T_q x$ is continuous, the set X_{mn} is closed and the hypothesis on the sequence $T_n x$ shows that X_{mn} is non-dense. Hence $X - X_1$ is of the first category in X.

Marcinkiewicz[3] constructed his example by taking X as the space of continuous functions, Y as the space of measurable functions, $\{y_m\}$ the set of polynomials with rational coefficients and $T_n x = [x(t + h_n) - x(t)]/h_n$. It is obvious that for each m there is an x such that $T_n x \to y_m$. Also in the ϵ-neighborhood of an arbitrary x there is an x_0 which is constructed like the Cantor function on each of a finite number of intervals (intervals upon which the oscillation of x is $< \epsilon$) such that $T_n x_0 \to 0$. Thus the set of x for which $T_n x \to y_m$ is everywhere dense in X. Since X is not of the first category, the existence of the continuous function $\varphi(t)$ described in the introduction follows from Theorem 1.

3. The construction of the sequence $f_n(t)$. A sequence of partitions of $(0, 1)$ is defined by

$$\delta_{m1} = (0 \leqq t \leqq 1/2^{m-1}),\quad \delta_{mi} = [(i-1)/2^{m-1} < t \leqq i/2^{m-1}],\quad (i = 2, 3, \cdots, 2^{m-1}).$$

A function $f(t)$ will be said to belong to the class X_m in case there is a positive number $\eta < 1/2^{m-1}$ such that

$$f(0) = 0,\quad f(t) = 0 \text{ for } [(i-1)/2^{m-1} < t \leqq (i-1)/2^{m-1} + \eta],$$
$$(i = 1, 2, \cdots, 2^{m-1}),$$

and $f(t)$ is constant on the remaining part of δ_{mi} $(i = 1, 2, \cdots, 2^{m-1})$.

LEMMA 1. *For every integer m the set $\sum_{n=m}^{\infty} X_n$ is dense in L [the space of real functions summable on $(0, 1)$].*

[3] This paragraph of exposition is inserted since the reader (without a critical examination of Marcinkiewicz' paper) may see no connection between Theorem 1 and the function $\varphi(t)$ described in the introduction.

Let $f(t)$ be an arbitrary summable function and $\epsilon > 0$. Then there is an integer $m' \geqq m$ and a function $g(t)$ which is constant on each of the intervals $\delta_{m'i}$, $i = 1, 2, \cdots, 2^{m'-1}$, such that

$$\| f - g \| = \int_0^1 | f(t) - g(t) | \, dt < \epsilon/2.$$

There is a positive number $\delta < 1$ such that $\int_e | g(t) | \, dt < \epsilon/2$ if $m(e) < \delta$. Thus if $\eta = \delta/2^{m'-1}$ and

$$h(0) = 0, \quad h(t) = 0 \text{ for } (i-1)/2^{m'-1} < t \leqq (i-1)/2^{m'-1} + \eta, \qquad (i = 1, \cdots, 2^{m'-1}),$$

$$h(t) = g(t) \text{ elsewhere on } (0, 1),$$

then $h(t)$ belongs to $X_{m'}$ and $\| f - h \| \leqq \| f - g \| + \| g - h \| < \epsilon$, which completes the proof.

Define the step functions $s_{mn}(t) = n2^{m-1}$ for t in the intervals

$$\delta(i, m, n) = [(i-1)/2^{m-1} < t \leqq (i-1)/2^{m-1} + 1/(n2^{m-1})], \quad (i = 1, \cdots, 2^{m-1}),$$

$$s_{mn}(t) = 0 \text{ elsewhere on } (0, 1).$$

Lemma 2. *For each m the set of functions $f(t)$ in L for which*

$$\lim_n \int_0^x s_{mn}(t) f(t) \, dt = 0 \qquad (0 \leqq x \leqq 1)$$

is everywhere dense in L.

Fix $f(t)$ in $\sum_{n=m}^{\infty} X_n$; then for some $m' \geqq m$ and some $\eta > 0$, $f(t) = 0$ on the intervals γ_j ($j = 1, \cdots, 2^{m'-1}$) where

$$\gamma_1 = (0 \leqq t \leqq \eta), \qquad \gamma_j = [(j-1)/2^{m'-1} < t \leqq (j-1)/2^{m'-1} + \eta].$$

Now $s_{mn}(t) = 0$ except on the intervals $\delta(i, m, n)$. Since $m' \geqq m$, the integer $j = 1 + (i-1)2^{m'-m}$ is no greater than $2^{m'-1}$, and for this j the interval γ_j contains $\delta(i, m, n)$, provided n is so large that $1/(n2^{m-1}) < \eta$. Thus for n sufficiently large, $s_{mn}(t)f(t) = 0$ for $0 \leqq t \leqq 1$. The conclusion follows from Lemma 1.

The set of functions each of which vanishes at the origin and is a rational constant on the rest of δ_{mi}, $i = 1, \cdots, 2^{m-1}$, for some $m = 1, 2, \cdots$ form a denumerable set everywhere dense in S, the space[4] of measurable functions. Let this set be ordered in any manner into the sequence $F_p(x)$.

[4] See Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 9.

LEMMA 3. *For every integer p there is an m_p and a function $f(t)$ in L such that*

$$\lim_n \int_0^x s_{m_p n}(t)\, f(t)dt = F_p(x) \qquad (0 \leqq x \leqq 1).$$

Suppose $F_p(x) = a_i$ for x in δ_{mi}, $i = 1, \cdots, 2^{m-1}$. Define $m_p = m$, $f(t) = a_1$ on δ_{m1}, $f(t) = a_i - a_{i-1}$ on δ_{mi}, $i = 2, 3, \cdots, 2^{m-1}$. Then for x in δ_{mi}

$$\int_0^x s_{mn}(t)\, f(t)dt = a_1 + a_2 - a_1 + \cdots + a_{i-1} - a_{i-2}$$
$$+ (a_i - a_{i-1}) \int_{\delta_{mi}(0,x)} s_{mn}(t)dt,$$

and since $\lim_n \int_{\delta_{mi}(0,x)} s_{mn}(t)dt = 1$, we have

$$\lim_n \int_0^x s_{mn}(t)\, f(t)dt = F_p(x) \qquad (0 \leqq x \leqq 1).$$

LEMMA 4. *For every integer p the set of functions $f(t)$ in L for which*

$$\lim_n \int_0^x s_{m_p n}(t)\, f(t)dt = F_p(x) \qquad (0 \leqq x \leqq 1)$$

is everywhere dense in L. (m_p is the integer of Lemma 3.)

This is an immediate corollary of Lemmas 2 and 3.

Now arrange the functions $s_{mn}(t)$ in a triangular array

$$\begin{array}{lll} s_{11}, & s_{12}, & s_{13}, \cdots \\ & s_{22}, & s_{23}, \cdots \\ & & s_{33}, \cdots, \end{array}$$

and define the sequence $f_1 = s_{11}, f_2 = s_{12}, f_3 = s_{22}, f_4 = s_{13}$, etc. The sequence $f_n(t)$ then has the property that $f_n(t) = 0$ except on a set whose measure approaches zero with $1/n$. Placing

$$T_n g = \int_0^x f_n(t)\, g(t)dt,$$

it is seen that $T_n g$ is a continuous function on L to S. In Theorem 1 take $X = L$, $Y = S$, $y_p = F_p$. From Lemma 4 it is seen that for each p there is an unbounded sequence of integers n_i such that $\lim T_{n_i} g = F_p$ for every g in a set everywhere dense in L. Thus Theorem 1 gives

THEOREM 2. *There is a sequence of positive step functions $f_n(t)$ such that $f_n(t) = 0$ except for a set whose measure approaches zero with $1/n$ and such that for every $g(t)$ in L except for those in a set of the first category the sequence*

$\int_0^x f_n(t)g(t)dt$ *has the following property. For every measurable function* $F(x)$ *there is a subsequence such that*

$$\lim_i \int_0^x f_{n_i}(t)\, g(t)dt = F(x) \quad \textit{almost uniformly}\,.$$

Since, for $p > 1$, L_p is of the first category in L, one might ask if the exceptional set contains L_p. This is not the case, for a reference to the argument shows that the same theorem holds if L is replaced by L_p (the sequence $f_n(t)$ remaining fixed). The following corollaries are considerably weaker than Theorem 2 itself.

COROLLARY 1. *If* f *is in* L *and* F *is in* S, *there is a sequence* f_n *in* L *such that*

$$\lim_n f_n(t) = f(t)\,, \qquad \lim_n \int_0^x f_n(t)dt = F(x) \quad \textit{almost uniformly}\,,$$

and another sequence g_n *in* L *such that*

$$\lim_n g_n(t) = f(t)\,, \qquad \overline{\lim_n} \int_0^x g_n(t)dt = +\infty\,, \qquad \underline{\lim_n} \int_0^x g_n(t)dt = -\infty$$

almost everywhere.

COROLLARY 2. *For* f *in* L *there is a set* S_f *dense in* S *such that for each* F *in* S_f *there is a sequence* f_n *in* L *such that*

$$\lim f_n(t) = f(t)\,, \qquad \lim_n \int_0^x f_n(t)dt = F(x) \quad \textit{everywhere}\,.^5$$

COROLLARY 3. *For* f *in* L *and* F *in* S *there is a sequence of absolutely continuous functions* F_n *such that*

$$F_n(x) \to F(x)\,, \qquad F_n'(t) \to f(t) \quad \textit{almost uniformly}\,.$$

BROWN UNIVERSITY.

[5] Corollary 2 is a corollary of Lemma 4.

THE PROBABILITY LIMIT THEOREM

By Arthur H. Copeland

The probability limit theorem is concerned with the asymptotic behavior of certain sequences of integrals, but this should not be all. It is intended to throw light upon the nature of physical measurements. Whether or not measurements do behave in the manner described, is, of course, not a mathematical question; but whether or not the assumption of such behavior implies inconsistency is a mathematical question. As I have pointed out in a previous paper, we can answer such questions of consistency by studying the behavior of certain infinite matrices.[1] These matrices are called variates.[2] It is possible to analyse any proof of the probability limit theorem in terms of the matrix theory of probability, but it is preferable to start with an entirely new proof which is based directly upon the properties of variates.[3]

We shall give a brief description of those properties of variates which will be used in this paper. A variate x is an infinite sequence of numbers, thus

$$x = x^{(1)}, x^{(2)}, x^{(3)}, \cdots, x^{(k)}, \cdots,$$

where $x^{(k)}$ is an arbitrary number. A variate is called a constant, or parameter, when all of its terms are the same. Thus $a = a, a, a, \cdots$ is a constant or parameter. The average of the first n terms of a variate x is denoted by $p_n(x)$, that is, $p_n(x) = \sum_{k=1}^{n} x^{(k)}/n$. We shall let $p(x) = \lim_{n\to\infty} p_n(x)$. Then $p(x)$ is called the expected value of x, or the first moment of x.

Let $x_1, x_2, \cdots, x_n$ be n variates and $f(s_1, s_2, \cdots, s_n)$ be a function of n variables. Then $f(x_1, x_2, \cdots, x_n)$ is a variate defined as follows:

$$f(x_1, x_2, \cdots, x_n) = f(x_1^{(1)}, x_2^{(1)}, \cdots, x_n^{(1)}), f(x_1^{(2)}, x_2^{(2)}, \cdots, x_n^{(2)}), \cdots.$$

For example, x^2 is a variate such that $p(x^2)$ is the second moment of x. Further, we shall mention two properties of the operator $p(\ )$. First, this operator is additive, i.e.,

$$p(x_1 + x_2 + \cdots + x_n) = p(x_1) + p(x_2) + \cdots + p(x_n).$$

Received by the Editors of the Annals of Mathematics, February 11, 1935; accepted by them, and later transferred to this journal; presented to the American Mathematical Society, September 11, 1931.

[1] See the author's article, reference (d), at the end of this paper.

[2] A variate is essentially the same as a Kollektiv. See reference (a).

[3] For further literature on the subject of the probability limit theorem, see Khintchine, reference (b). That paper includes an extensive bibliography.

Second, if a is any constant, $p(a) = a$ and $p(ax) = a \cdot p(x)$. Thus $x' = x - p(x)$ is a variate such that $p(x') = 0$.

Let $\varphi_I(s)$ denote the fundamental function of the interval I: $a \leqq y < b$. Then $\varphi_I(s)$ is equal to 1, if s lies in I, and is equal to 0 otherwise. The variate $\varphi_I(x)$ can be interpreted as an event which succeeds on its k-th trial if and only if the k-th term of x lies in I, success being denoted by 1 and failure by 0. Then $\sum_{k=1}^{n} \varphi_I(x^{(k)})$ is equal to the number of 1's in the first n terms of $\varphi_I(x)$. Thus $p_n[\varphi_I(x)]$ is the success ratio and $p[\varphi_I(x)]$ is the probability of the event $\varphi_I(x)$. Let I_s and I_s' denote respectively the intervals $-\infty < y \leqq s$ and $-\infty < y < s$, and let $p[\varphi_{I_s}(x)] = F(s + 0)$, $p[\varphi_{I_s'}(x)] = F(s - 0)$, and $F(s) = [F(s + 0) + F(s - 0)]/2$. Then $F(s)$ is called the distribution function of the variate x. It will be observed that $F(s + 0), F(s - 0), F(s + 0) - F(s - 0)$ are, respectively, the probabilities that a term of x will be $\leqq s$, $< s$, $= s$.

We shall define the dependence and independence of variates in terms of the fundamental function. A set $x_1, x_2, \cdots, x_n, \cdots$ (finite or infinite) is said to be a set of dependent variates if there exists a finite subset $x_{n_1}, x_{n_2}, \cdots, x_{n_k}$ and a corresponding set of intervals $I_1, I_2, \cdots, I_k$ such that

$$p[\varphi_{I_1}(x_{n_1}) \cdot \varphi_{I_2}(x_{n_2}) \cdots \varphi_{I_k}(x_{n_k})] \neq p[\varphi_{I_1}(x_{n_1})] \cdot p[\varphi_{I_2}(x_{n_2})] \cdots p[\varphi_{I_k}(x_{n_k})].$$

Variates which are not dependent are said to be independent. It will be observed that the independence of the variates of a given set implies the existence of their distribution functions.

We shall consider an infinite set of independent variates $x_1, x_2, \cdots, x_k, \cdots$ such that $p(x_k) = 0$. Let $X_n = (x_1 + x_2 + \cdots + x_n)/B_n$, where $B_n = b_1^2 + b_2^2 + \cdots + b_n^2$ and $b_k^2 = p(x_k^2)$. The probability limit theorem, which we shall now state, is concerned with the distribution function of the variate X_n.

THEOREM. *If the variates $x_1, x_2, \cdots, x_k, \cdots$ are independent and if, given any positive number ϵ, there exist two numbers a and N such that $p[\varphi_{|s| \geqq a}(x_k) \cdot x_k^2] < b_k^2 \cdot \epsilon$ for every k, and such that $b_k^2/B_n^2 < \epsilon$ for every k less than or equal to n whenever $n \geqq N$, then*

$$\lim_{n \to \infty} p[\varphi_{I_s}(X_n)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-t^2/2}\, dt.$$

There is no essential restriction involved in the condition $p(x_k) = 0$, since if this condition does not hold, we can set $x_k' = x_k - p(x_k)$ and apply the theorem to the variates $x_1', x_2', \cdots$.

We shall give an outline of the method of proof of the limit theorem. Equations (1) and (2), which are derived in this outline, will be used in the formal proof. The proof is accomplished by the aid of the characteristic function. If x is a variate and t is a parameter, the characteristic function of x is the expression $p(e^{ixt})$. The characteristic function of a variate always exists whenever the distribution function exists.[4] The variate X_n consists of a sum of

[4] See (d), pp. 545–547.

variates and $e^{ix_n t}$ consists of a product of variates. Since these variates are independent, the expected value of their product is equal to the product of their expected values. Hence

$$(1) \qquad p(e^{ix_n t}) = \prod_{k=1}^{n} p(e^{ix_k t/B_n}).$$

By means of this equation, it will be proved (Lemma 1) that $\lim\limits_{n\to\infty} p(e^{ix_n t}) = e^{-t^2/2}$. Next we must obtain a method of computing the distribution function in terms of the characteristic function (Lemma 2). We have the equation

$$\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{e^{it} - e^{ixt}\cdot e^{-ist}}{t}\,dt = \begin{cases} 1 & \text{if } x < s, \\ 1/2 & \text{if } x = s, \\ 0 & \text{if } x > s. \end{cases}$$

If x is a variate, the integral is a variate such that there is a probability $F(s-0)$ that one of its terms will be equal to 1 and a probability $F(s+0) - F(s-0)$ that one of its terms will be equal to 1/2. Hence

$$(2) \quad p\left[\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{e^{it} - e^{ixt}\cdot e^{-ist}}{it}\,dt\right] = F(s-0) + \frac{F(s+0) - F(s-0)}{2} = F(s).$$

We shall prove that we can interchange the order of the operation of $p(\ \ )$ with that of integration. When the operator $p(\ \ )$ is applied to the integrand, it affects only the expression e^{ixt}, since the remaining portions of the integrand depend on parameters and not on x. Thus[5]

$$\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{e^{it} - p(e^{ixt})\cdot e^{-ist}}{t}\,dt = F(s).$$

Finally, (Lemma 3)

$$\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{e^{it} - e^{-t^2/2}\cdot e^{-ist}}{t}\,dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{s} e^{-t^2/2}\,dt.$$

We shall now turn to the formal proof of the three lemmas and the theorem.

LEMMA 1. *Under the hypotheses of the limit theorem*

$$|p(e^{ix_n t}) - e^{-t^2/2}| < e^{-t^2/2}\cdot\epsilon_n\cdot t^2,$$

where $\lim\limits_{n\to\infty} \epsilon_n = 0$.

Let $e^{it} = 1 + it - t^2[1 + R(t)]/2$. Then $R(t) = 2(1 + it - e^{it})/t^2 - 1$. The function $R(t)$ is analytic and bounded for all real values of t. Furthermore, $R(0) = 0$ and $R(\pm\infty) = -1$. Thus

$$p(e^{ix_k t/B_n}) = 1 - \frac{b_k^2 t^2}{2\cdot B_n^2}(1 + r_{n,k}),$$

[5] Paul Lévy uses a slightly different integral for this purpose. See (c).

where

$$\frac{b_k^2t^2}{2\cdot B_n^2}\cdot r_{n,k} = p\left[\frac{x_k^2t^2}{2\cdot B_n^2}\cdot R\left(\frac{x_kt}{2B_n}\right)\right].$$

Moreover

$$\frac{b_k^2t^2}{2\cdot B_n}\cdot | r_{n,k} | \leqq p\left[\frac{x_k^2t^2}{2\cdot B_n^2}\cdot \left|R\left(\frac{x_kt}{2B_n}\right)\right|\right]$$
$$= p\left[\frac{x_k^2t^2}{2\cdot B_n^2}\cdot \left|R\left(\frac{x_kt}{2B_n}\right)\right|\cdot \varphi_{|s|\geqq a}(x_k)\right] + p\left[\frac{x_k^2t^2}{2\cdot B_n^2}\left|R\left(\frac{x_kt}{2B_n}\right)\right|\cdot \varphi_{|s|<a}(x_k)\right].$$

By hypothesis, given a positive number δ, there exists a number a such that

$$p\left[\frac{x_k^2t^2}{2\cdot B_n^2}\cdot \left|R\left(\frac{x_kt}{2B_n}\right)\right|\cdot \varphi_{|s|\geqq a}(x_k)\right] < \frac{\delta b_k^2t^2}{4B_n^2}.$$

There exists a number m such that $| R(st/B_n) | < \delta/4$ if $| st | /B_n \leqq m$, that is, if $| s | \leqq m\cdot B_n/ | t |$. Let h be an arbitrary positive number. There exists a number N such that $m\cdot B_n/ |t| \geqq a$ if $|t| \leqq h$ and $n \geqq N$. Therefore

$$p\left[\frac{x_k^2t^2}{2\cdot B_n^2}\cdot \left|R\left(\frac{x_kt}{2B_n}\right)\right|\cdot \varphi_{|s|<a}(x)\right] < \frac{\delta b_k^2t^2}{4B_n^2},$$

and hence $| r_{n,k} | < \delta$ whenever $| t | < h$ and $n \geqq N$. We have the equation

$$\log\, p(e^{ix_kt/B_n}) = -\frac{b_k^2t^2}{2\cdot B_n^2}[1 + \rho_{n,k}],$$

where

$$\frac{b_k^2t^2}{2\cdot B_n^2}\rho_{n,k} = \frac{b_k^2t^2}{2\cdot B_n^2}r_{n,k} + \left[\frac{b_k^2t^2}{2B_n^2}\cdot(1 + r_{n,k})\right]^2\Big/ 2 + \cdots.$$

Hence by choosing n sufficiently large, all of the numbers $| \rho_{n,k} |$ such that $k \leqq n$ can be made arbitrarily small. From equation (1) it follows that $\log p(e^{i\chi_nt}) = - t^2(1 + \rho_n)/2$ or $p(e^{i\chi_nt}) = e^{-t^2(1+\rho_n)/2}$, where $| \rho_n |$ is less than or equal to the largest of the numbers $| \rho_{n,k} |$. Therefore $| p(e^{i\chi_nt}) - e^{-t^2/2} | < e^{-t^2/2}\cdot \epsilon_n\cdot t^2$, where $\lim_{n\to\infty} \epsilon_n = 0$.

LEMMA 2. *If x is a variate whose distribution function $F(s)$ exists, and if* $F(s, h) = \dfrac{1}{2\pi i}\displaystyle\int_{-h}^{+h}\frac{e^{it} - p(e^{ixt})\cdot e^{-ist}}{t}\,dt$ *and* $\epsilon = 1/\sqrt{h}$, *then*

$$| F(s) - F(s, h) | < 2\epsilon + [F(s + \epsilon) - F(s - \epsilon)]/2 .$$

The function $(e^{it} - e^{iut}\cdot e^{-ist})/t$ is bounded in the entire u, t-plane and is continuous in both variables. Thus as t becomes infinite, $p_n[(e^{it} - e^{ixt}\cdot e^{-ist})/t]$ con-

verges uniformly for $|t| \leqq h$, and[6] hence $p\left(\frac{1}{2\pi i}\int_{-h}^{+h}\frac{e^{it}-e^{ixt}\cdot e^{-ist}}{t}\,dt\right) = F(s, h)$. From equation (2) it follows that $F(s) - F(s, h) = p(v)$, where

$$v = \frac{1}{2\pi i}\int_{|t|\geqq h}\frac{e^{it}-e^{ixt}\,e^{-ist}}{t}\,dt$$
$$= \frac{1}{\pi}\int_h^\infty \frac{\sin t}{t}\,dt - \frac{\operatorname{sgn}(x-s)}{\pi}\int_{|x-s|\cdot h}^\infty \frac{\sin t}{t}\,dt\,,$$

and where $\operatorname{sgn}(x - s) = 1, 0, -1$ according as $x >, =, < s$.

We have the inequalities (a) $\left|\frac{1}{\pi}\int_h^\infty \frac{\sin t}{t}\,dt\right| < 1/h < \epsilon$ if $1 < h$, (b) $\left|\frac{1}{\pi}\int_{|x-s|\cdot h}^\infty \frac{\sin t}{t}\,dt\right| < \epsilon$ if $|x - s| \geqq \epsilon$, (c) $\left|\frac{1}{\pi}\int_{|x-s|\cdot h}^\infty \frac{\sin t}{t}\,dt\right| \leqq 1/2$ for all values of x, s and h. Thus v is a variate such that $|v^{(k)}| < 2\epsilon$ if $|x^{(k)} - s| \geqq \epsilon$ and $|v^{(k)}| < 1/2$ otherwise. The probability that $|x^{(k)} - s| < \epsilon$ is less than or equal to $F(s + \epsilon) - F(s - \epsilon)$. Therefore

$$|F(s) - F(s, h)| < 2\epsilon + [F(s + \epsilon) - F(s - \epsilon)]/2\,.$$

Lemma 3. *If* $\Phi(s) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{s} e^{-t^2/2}\,dt$, $\Phi(s, h) = \frac{1}{2\pi i}\int_{-h}^{+h}\frac{e^{it} - e^{-t^2/2}\,e^{-ist}}{t}\,dt$ *and* $\epsilon = 1/\sqrt{h}$, *then* $|\Phi(s) - \Phi(s, h)| < 3\epsilon$.

Let u be a variate which has the distribution function $\Phi(s)$. Then[7]

$$p(e^{iut}) = \int_{-\infty}^{+\infty} e^{ist}\,d\Phi(s) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{ist}\,e^{-s^2/2}\,ds = \frac{e^{-t^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-(s-it)^2/2}\,ds\,.$$

Since there are no discontinuities of the function $e^{-(s-it)^2/2}$ (except at infinity) in the region bounded by the real axis and the line $s = it$, it follows that

$$p(e^{iut}) = e^{-t^2/2}\cdot\Phi(+\infty) = e^{-t^2/2}\,.$$

Hence by Lemma 2, $|\Phi(s) - \Phi(s, h)| < 2\epsilon + [\Phi(s + \epsilon) - \Phi(s - \epsilon)]/2$. Since $\Phi(s)$ satisfies the Lipschitz condition $\Phi(s + \epsilon) - \Phi(s) < \epsilon$, we get

$$|\Phi(s) - \Phi(s, h)| < 3\epsilon\,.$$

Let $\Phi_n(s) = p[\varphi_{I_s}(X_n)]$ and $\Phi_n(s, h) = \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{e^{it} - e^{iX_n t}\,e^{-ist}}{t}\,dt$. From Lemma 1 we conclude that there exists a number N such that

$$|\Phi_n(s, h) - \Phi(s, h)| < \epsilon \text{ whenever } n \geqq N\,. \tag{3}$$

From Lemmas 2 and 3 we obtain the inequalities

[6] See (d), Theorem 3.
[7] See (d), Theorem 2.

$$
(4) \qquad \begin{aligned} | \Phi_n(s) - \Phi_n(s, h) | &< 2\epsilon + [\Phi_n(s + \epsilon) - \Phi_n(s - \epsilon)]/2 \, , \\ | \Phi(s, h) - \Phi(s) | &< 3\epsilon \, . \end{aligned}
$$

Combining (3) and (4) we get

$$
(5) \qquad | \Phi_n(s) - \Phi(s) | < 6\epsilon + [\Phi_n(s + \epsilon) - \Phi(s - \epsilon)]/2 \, .
$$

This inequality can be replaced by

$$
(6) \qquad \begin{aligned} \Phi_n(s - \epsilon) - 6\epsilon - [\Phi_n(s + \epsilon) - \Phi_n(s)]/2 &< \Phi(s) \\ &< \Phi_n(s + \epsilon) + 6\epsilon + [\Phi_n(s) - \Phi_n(s - \epsilon)]/2 \, . \end{aligned}
$$

Replacing s by $s + 2\epsilon$ we see that

$$
(7) \qquad \begin{aligned} \Phi_n(s) < \Phi(s) + 8\epsilon \text{ whenever } & \Phi_n(s + \epsilon) - \Phi_n(s) \\ & \geqq [\Phi_n(s + 3\epsilon) - \Phi_n(s + 2\epsilon)]/2 \, . \end{aligned}
$$

We shall consider separately the cases $\Phi_n(s + \epsilon) - \Phi_n(s) > 2\epsilon$ and $\Phi_n(s + \epsilon) - \Phi_n(s) \leqq 2\epsilon$. Let us assume that $\Phi_n(s) \geqq \Phi(s) + 8\epsilon$ and $\Phi_n(s + \epsilon) - \Phi_n(s) > 2\epsilon$. Then $\Phi_n(s + 2\epsilon) \geqq \Phi_n(s + \epsilon) > \Phi(s) + 10\epsilon > \Phi(s + 2\epsilon) + 8\epsilon$, and by condition (7), $\Phi_n(s + 3\epsilon) - \Phi_n(s + 2\epsilon) > 4\epsilon > 2\epsilon$. By induction $\Phi_n(s + 2k\epsilon) > \Phi(s) + 8\epsilon + 2k\epsilon$, or $\lim_{k \to \infty} \Phi_n(s + 2k\epsilon) = + \infty$.

Since this is impossible, we have

$$
(8) \qquad \Phi_n(s - \epsilon) \leqq \Phi_n(s) < \Phi(s) + 8\epsilon \text{ whenever } \Phi_n(s + \epsilon) - \Phi_n(s) > 2\epsilon \, .
$$

But it follows from (6) that

$$
(9) \qquad \Phi_n(s - \epsilon) < \Phi(s) + 7\epsilon \text{ whenever } \Phi_n(s + \epsilon) - \Phi_n(s) \leqq 2\epsilon \, .
$$

Replacing s by $s + \epsilon$ in (8) and (9) we get $\Phi_n(s) < \Phi(s) + 9\epsilon$, and similarly, $\Phi(s) - 9\epsilon < \Phi_n(s)$. Therefore $\lim_{n \to \infty} \Phi_n(s) = \Phi(s)$.

This proof of the probability limit theorem is based on assumptions which are expressed in terms of the properties of independent variates. Hence there is no inconsistency in the assumption that physical measurements behave in the manner described by the theorem.

References

(a) R. von Mises, *Grundlagen der Wahrscheinlichkeitsrechnung*, Math. Zeitschr., vol. 5 (1919), pp. 52–99.

(b) A. Khintchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik und ihrer Grenzgebiete.

(c) Paul Lévy, *Calcul des Probabilités*, Paris.

(d) A. H. Copeland, *A matrix theory of measurement*, Math. Zeitschr., vol. 37 (1933), pp. 542–555.

The University of Michigan.

APPLICABILITY WITH PRESERVATION OF BOTH CURVATURES

By W. C. Graustein

1. **Introduction.** The determination of conditions necessary and sufficient that there exist a surface applicable to a given surface with preservation of both the total and mean curvatures constitutes a problem of classical differential geometry which has received no little attention.[1] In this paper, various new conditions, all in invariantive form, are found. The map of a surface satisfying these conditions on a surface applicable to it in the manner described is studied in some detail and is shown to have many interesting geometrical properties.

The treatment is by means of the invariant methods recently exploited by the author.[2] These methods are particularly advantageous in the present problem, in that they naturally disclose facts which otherwise might remain undiscovered or prove complicated to establish.

2. **Necessary and sufficient conditions.** Let there be given a surface[3] $S: x_i = x_i(u, v)$, $i = 1, 2, 3$, and assume that there exists a surface $\bar{S}: \bar{x}_i = \bar{x}_i(u, v)$, $i = 1, 2, 3$, which is applicable to S so that both curvatures are preserved. Then any surface $\bar{S}^*$ which is symmetric to $\bar{S}$ is also applicable to S with preservation of both curvatures. Inasmuch as the sign of the mean curvature of a surface depends on the orientation of the directed normal, it follows that the normals to $\bar{S}$ and $\bar{S}^*$ must be so directed that corresponding directions of rotation about corresponding points have, with reference to these directed normals, opposite senses. Hence, for just one of the surfaces $\bar{S}$, $\bar{S}^*$, the map of the surface on S has the property that corresponding directions of rotation about corresponding points are the same. Without loss of generality we may assume that this is the surface $\bar{S}$; the surface $\bar{S}^*$ we then exclude completely, since it is readily obtainable from $\bar{S}$. In other words, we assume that the normals of two surfaces which are applicable to one another with preservation of both curvatures are so directed that (without invalidating the equality of the mean curvatures) corresponding directions of rotation about corresponding points, referred to these directed normals, have the same sense.

Received January 16, 1936.

[1] For references to the literature see Graustein, *Applicability with preservation of both curvatures*, Bull. Amer. Math. Soc., vol. 30 (1924), pp. 19–23. This paper will be referred to later as "Paper A".

[2] *Méthodes invariantes dans la géométrie infinitésimale des surfaces*, Mémoires de l'Académie Royale de Belgique (Classe des Sciences), (2), vol. 11 (1929); *Invariant methods in classical differential geometry*, Bull. Amer. Math. Soc., vol. 36 (1930), pp. 489–521. These papers will be referred to respectively as "B.M." and "I.M".

[3] It is assumed that all functions are real, single-valued, and analytic in a certain domain of the real variables u, v.

From the fact that the expressions for the total and mean curvatures of a surface in terms of the principal normal curvatures are symmetric in the latter curvatures we draw the following conclusion.

LEMMA. *A necessary and sufficient condition that two surfaces be applicable with preservation of both curvatures is that they be applicable* (a) *so that the principal normal curvatures are preserved, or* (b) *so that in the neighborhoods of corresponding points they are congruent through infinitesimals of the second order.*

We return to our given surfaces S and $\bar{S}$. Let $1/r_1$, $1/r_2$ be the common principal normal curvatures and denote the corresponding families of lines of curvature, individually and collectively, by C_1, C_2, in the case of S, and by $\bar{C}_1, \bar{C}_2$, in the case of $\bar{S}$. Let C, C' symbolize the two families of curves on S corresponding respectively to the families of lines of curvature $\bar{C}_1, \bar{C}_2$ on $\bar{S}$ and let $\bar{C}, \bar{C}'$ represent the two families of curves on $\bar{S}$ which correspond respectively to the families C_1, C_2 of lines of curvature on S.

Let P be an arbitrary point of S and direct the curves C_1, C_2, C, C' through P (a) so that the direction of rotation from the directed curve C_1 to the directed curve C_2 and that from the directed curve C to the directed curve C' are both positive, and (b) so that, for convenience, the smallest non-negative directed angle α from the curves C_1, C_2 to the curves C, C', respectively, is less than π: $0 \leqq \alpha < \pi$.

To the curves $\bar{C}, \bar{C}', \bar{C}_1, \bar{C}_2$ through the point $\bar{P}$ of $\bar{S}$ corresponding to the point P of S we give the directions which correspond, by the map of S on $\bar{S}$, to the directions assigned to C_1, C_2, C, C'. Then the direction of rotation from the directed curve $\bar{C}_1$ to the directed curve $\bar{C}_2$ and that from the directed curve $\bar{C}$ to the directed curve $\bar{C}'$ are both positive, and the rotation about $\bar{P}$ through the angle $-\alpha$ carries the positive directions of the curves $\bar{C}_1, \bar{C}_2$ into those of the curves $\bar{C}, \bar{C}'$, respectively.

It follows, from this last fact, that, if $1/\bar{r}$, $1/\bar{r}'$ are the normal curvatures, and $1/\bar{\tau}$, $1/\bar{\tau}'$ $(= -1/\bar{\tau})$ are the geodesic torsions of the curves $\bar{C}, \bar{C}'$, respectively, we have

$$\frac{1}{\bar{r}} = \frac{\cos^2\alpha}{r_1} + \frac{\sin^2\alpha}{r_2}, \qquad \frac{1}{\bar{r}'} = \frac{\sin^2\alpha}{r_1} + \frac{\cos^2\alpha}{r_2}, \qquad \frac{2}{\bar{\tau}} = \left(\frac{1}{r_2} - \frac{1}{r_1}\right)\sin 2\alpha .$$

In terms of the quantities

$$K' = \frac{1}{r_1} + \frac{1}{r_2}, \qquad L = \frac{1}{r_1} - \frac{1}{r_2},$$

$$\bar{K}' = \frac{1}{\bar{r}} + \frac{1}{\bar{r}'}, \qquad \bar{L} = \frac{1}{\bar{r}} - \frac{1}{\bar{r}'},$$

these equations take the more convenient forms

$$(1) \qquad \bar{K}' = K', \qquad \bar{L} = L\cos 2\alpha, \qquad \frac{2}{\bar{\tau}} = -L\sin 2\alpha .$$

The Codazzi equations for S, expressed in terms of the directional derivatives $\partial/\partial s_1$, $\partial/\partial s_2$ in the positive directions of the lines of curvature C_1, C_2, may be written[4]

$$\frac{\partial K'}{\partial s_1} = \frac{2}{\rho_2} L + \frac{\partial L}{\partial s_1}, \qquad \frac{\partial K'}{\partial s_2} = \frac{2}{\rho_1} L - \frac{\partial L}{\partial s_2}, \tag{2}$$

where $1/\rho_1$, $1/\rho_2$ are respectively the geodesic curvatures of the directed curves C_1, C_2.

Since geodesic curvature and directional differentiation are preserved by the map of S on $\bar{S}$, the Codazzi equations of $\bar{S}$, referred to the orthogonal system of curves[5] $\bar{C}$, $\bar{C}'$, are reducible to the forms

$$\begin{aligned} &\frac{\partial \bar{K}'}{\partial s_1} - \frac{\partial \bar{L}}{\partial s_1} - \frac{2}{\rho_2}\bar{L} + \frac{\partial}{\partial s_2}\left(\frac{2}{\bar{\tau}}\right) - \frac{2}{\rho_1}\frac{2}{\bar{\tau}} = 0, \\ &\frac{\partial \bar{K}'}{\partial s_2} + \frac{\partial \bar{L}}{\partial s_2} - \frac{2}{\rho_1}\bar{L} + \frac{\partial}{\partial s_1}\left(\frac{2}{\bar{\tau}}\right) + \frac{2}{\rho_2}\frac{2}{\bar{\tau}} = 0. \end{aligned} \tag{3}$$

When we substitute for $\bar{K}'$, $\bar{L}$, $1/\bar{\tau}$ in (3) their values from (1), and solve the resulting equations for $\partial\alpha/\partial s_1$, $\partial\alpha/\partial s_2$, we obtain, by virtue of (2), the equations

$$\frac{\partial z}{\partial s_1} = Mz - N, \qquad \frac{\partial z}{\partial s_2} = Nz + M, \tag{4}$$

where

$$z = \cot\alpha$$

and[6]

$$M = \frac{1}{L}\frac{\partial K'}{\partial s_1} = \frac{\partial \log L}{\partial s_1} + \frac{2}{\rho_2}, \qquad N = -\frac{1}{L}\frac{\partial K'}{\partial s_2} = \frac{\partial \log L}{\partial s_2} - \frac{2}{\rho_1}. \tag{5}$$

Equations (4) constitute necessary conditions on the surface S and the angle α that there exist a surface $\bar{S}$ applicable to S with preservation of both curvatures. These conditions are also sufficient. For, if equations (4) are fulfilled for S by a specific α, the values of $\bar{K}'$, $\bar{L}$, $2/\bar{\tau}$ obtained from (1) for this α satisfy (3), a surface $\bar{S}$ with the desired properties is uniquely determined to within its position in space, and the angle α has the proper interpretation[7] on both S and $\bar{S}$.

Employing the relation

$$\frac{\nabla}{\nabla s_2}\frac{\partial z}{\partial s_1} = \frac{\nabla}{\nabla s_1}\frac{\partial z}{\partial s_2},$$

[4] See B.M., p. 39.

[5] See B.M., p. 53; I.M., p. 508.

[6] The assumption that $L \neq 0$ is easily justified; a sphere evidently admits no surface applicable to it with preservation of both curvatures.

[7] See B.M., pp. 53, 54.

where $\nabla/\nabla s_1$, $\nabla/\nabla s_2$ are the modified directional derivatives[8] in the positive directions of the lines of curvature on S, we obtain, as the condition of compatibility of equations (4),

$$Pz - Q = 0\,, \tag{6}$$

where

$$P = \frac{\nabla M}{\nabla s_2} - \frac{\nabla N}{\nabla s_1}\,, \qquad Q = \frac{\nabla M}{\nabla s_1} + \frac{\nabla N}{\nabla s_2} - M^2 - N^2\,. \tag{7a}$$

Substituting for M, N in (7a) properly chosen values from (5) and making use of the following expressions for the differential parameters[9] $\Delta_1\varphi$, $\Delta_2\varphi$ and the total curvature[10] K,

$$\Delta_1\varphi = \left(\frac{\partial\varphi}{\partial s_1}\right)^2 + \left(\frac{\partial\varphi}{\partial s_2}\right)^2, \qquad \Delta_2\varphi = \frac{\nabla}{\nabla s_1}\frac{\partial\varphi}{\partial s_1} + \frac{\nabla}{\nabla s_2}\frac{\partial\varphi}{\partial s_2}\,,$$

$$K = \frac{\nabla}{\nabla s_2}\left(\frac{1}{\rho_1}\right) - \frac{\nabla}{\nabla s_1}\left(\frac{1}{\rho_2}\right),$$

we find, as alternative values of P, Q,

$$P = 2\left[\frac{\nabla}{\nabla s_1}\left(\frac{1}{\rho_1}\right) + \frac{\nabla}{\nabla s_2}\left(\frac{1}{\rho_2}\right)\right], \qquad Q = \Delta_2 \log L - 2\,K - \frac{1}{L^2}\Delta_1 K'\,. \tag{7b}$$

From the first of these formulas it follows that the vanishing of P is a necessary and sufficient condition that S be an isometric surface.[11] Hence we may pass to the following conclusions.

THEOREM 1. *There are ∞^1 surfaces S applicable to a given surface S with preservation of both curvatures if S is an isometric surface and*

$$L^2\Delta_2 \log L = 2KL^2 + \Delta_1 K'\,.$$

There is a unique surface $\bar{S}$ if S is not isometric and $z = Q/P$ satisfies equations (4). *Otherwise, no surface $\bar{S}$ exists.*[12]

3. Surfaces admitting ∞^1 applicable surfaces. If S is an isometric surface, $P = 0$ and, by (7a), $\nabla M/\nabla s_2 = \nabla N/\nabla s_1$. Hence if $ds_1 = A_2du + B_2dv$ and $ds_2 = A_1du + B_1dv$ are the differentials of arc of the directed lines of curvature C_1, C_2 on S, $Mds_1 + Nds_2$ is the exact differential of a function.[13] Writing

[8] See B.M., pp. 57, 78; I.M., p. 500.
[9] See B.M., pp. 60, 61; I.M., pp. 517, 518.
[10] See B.M., p. 58; I.M., p. 511.
[11] See B.M., pp. 55, 59; I.M., p. 513.
[12] For the corresponding theorem in non-invariant form, see Paper A, p. 23.
[13] See B.M., p. 58; I.M., p. 500.

this function in the form $-\log \varphi$, and using the second values of M and N given in (5), we find that

$$\varphi = \frac{1}{L} e^{-2S}, \quad \text{where} \quad S = \int \left(\frac{ds_1}{\rho_2} - \frac{ds_2}{\rho_1}\right). \tag{8}$$

Furthermore, since $M = -\partial \log \varphi/\partial s_1$, $N = -\partial \log \varphi/\partial s_2$, we have from (7a) the fact that

$$Q = -\frac{1}{\varphi} \Delta_2 \varphi .$$

Thus we have shown that S admits ∞^1 surfaces applicable to it with preservation of both curvatures if and only if it is an isometric surface and $\Delta_2\varphi = 0$. According to (5), $M = 0$ and $N = 0$, that is, φ is constant, when and only when S is of constant mean curvature. But, by (2), every surface for which K' is constant is isometric. Hence our result may be stated as follows.

THEOREM 2. *A necessary and sufficient condition that there exist ∞^1 surfaces applicable to the surface S so that both curvatures are preserved is that S be a surface of constant mean curvature or an isometric surface of variable mean curvature on which the curves $\varphi =$ const., where φ is given by (8), form an isometric family with φ as an isometric parameter.*

The angle α. When the values of M and N in terms of φ are substituted in (4), these equations become

$$\frac{\partial(\varphi z)}{\partial s_1} = \frac{\partial \varphi}{\partial s_2}, \qquad \frac{\partial(\varphi z)}{\partial s_2} = -\frac{\partial \varphi}{\partial s_1}.$$

Hence if ψ is a function to which φ is "conjugate",[14] the general solution of (4) is

$$z = \frac{\psi + k}{\varphi}, \tag{9}$$

where k is an arbitrary constant.

If there is a constant angle α which satisfies (4), $M = 0$, $N = 0$ and S has constant mean curvature. Conversely, if K' is constant, φ is constant, ψ is constant, and hence, by (9), α is constant.

THEOREM 3. *The angle α from the lines of curvature on S to the curves on S corresponding to the lines of curvature on $\bar{S}$ is constant only when the surfaces have constant mean curvature.*

It follows that only in this case do the lines of curvature on S correspond to the lines of curvature on a surface $\bar{S}$. If the surface $\bar{S}$ for which $\alpha = \alpha$ is denoted by $\bar{S}_\alpha$, $0 \leqq \alpha < \pi$, this surface is $\bar{S}_{\pi/2}$ and the lines of curvature $\bar{C}_1$, $\bar{C}_2$ on it correspond respectively to the lines of curvature C_2, C_1 on $\bar{S}_0 = S$. More generally,

[14] The functions ψ and φ are actually conjugate functions of isometric parameters for the lines of curvature on S, as is readily proved by referring S to these parameters. See also Paper A, p. 21.

the lines of curvature on every pair of surfaces $\bar{S}_\alpha$, $\bar{S}_{\alpha+\pi/2}$, $0 \leqq \alpha < \pi/2$, correspond.[15]

The counterpart of Theorem 3 in the general case is the following.

THEOREM 4. *If S is a surface which admits ∞^1 surfaces applicable to it with preservation of both curvatures and four of these surfaces are chosen, a specific cross-ratio of the four directions at a point P on S, which correspond, at the image points of P on the four surfaces, to the four principal directions associated with a chosen principal normal curvature, is the same at every point P on S.*

The result follows directly from the fact that equation (9) defines a linear transformation of k into z. In particular, if we take as the four surfaces $\bar{S}_\infty = S$, $\bar{S}_0$, $\bar{S}_1$, $\bar{S}_k$, where $\bar{S}_k$ is the surface $\bar{S}$ for which $k = k$, and denote the four corresponding directions at P by d_∞, d_0, d_1, d_k, we find that the cross-ratio $(d_\infty d_0, d_1 d_k)$ has the value k and thus obtain a geometric interpretation of k as the projective coördinate of the direction d_k referred to d_∞, d_0, d_1 as basic directions.

4. **Properties of the map. Primary and secondary orthogonal systems.** It follows from the lemma of §2 that two families of curves lying respectively on two surfaces which are applicable with preservation of both curvatures and making on these surfaces the same directed angle with the lines of curvature associated with a specific principal normal curvature have the same normal curvature and geodesic torsion. For the normal curvature and geodesic torsion are uniquely determined by the principal normal curvatures and the directed angle in question.

We recall also from §2 that if a family of curves $\mathfrak{C}$ on S has the slope angle θ with respect to the directed lines of curvature C_1, the corresponding family of curves $\bar{\mathfrak{C}}$ on $\bar{S}$ has the slope angle $\theta - \alpha$ with respect to the directed lines of curvature $\bar{C}_1$. Hence if $1/\mathfrak{r}$, $1/\mathfrak{t}$ and $1/\bar{\mathfrak{r}}$, $1/\bar{\mathfrak{t}}$ are respectively the normal curvature and geodesic torsion of the curves $\mathfrak{C}$ and $\bar{\mathfrak{C}}$, we have

$$(10)\qquad \begin{aligned} \frac{1}{\mathfrak{r}} &= \frac{1}{2}(K' + L\cos 2\theta), & \frac{1}{\mathfrak{t}} &= \frac{1}{2}L\sin 2\theta, \\ \frac{1}{\bar{\mathfrak{r}}} &= \frac{1}{2}[K' + L\cos 2(\theta - \alpha)], & \frac{1}{\bar{\mathfrak{t}}} &= \frac{1}{2}L\sin 2(\theta - \alpha). \end{aligned}$$

If $1/\bar{\mathfrak{r}} = 1/\mathfrak{r}$ and $1/\bar{\mathfrak{t}} = 1/\mathfrak{t}$, it follows that $\alpha = 0$. In other words:

THEOREM 5. *If two corresponding families of curves lying respectively on two surfaces which are mapped isometrically with preservation of both curvatures have the same normal curvature and geodesic torsion, the surfaces are congruent.*

If $1/\mathfrak{r} = 1/\bar{\mathfrak{r}}$ and $\alpha \neq 0$, it follows that $\theta = \alpha/2$ or $\theta = \alpha/2 + \pi/2 \pmod{\pi}$. Furthermore, $1/\mathfrak{t}$ for $\theta = \alpha/2$ is equal to $1/\bar{\mathfrak{t}}$ for $\theta = \alpha/2 + \pi/2$, and $1/\mathfrak{t}$ for $\theta = \alpha/2 + \pi/2$ is equal to $1/\bar{\mathfrak{t}}$ for $\theta = \alpha/2$. Hence we have the proposition.

[15] If $K' = 0$, a closed continuum of surfaces S_α consists of a family of associate minimal surfaces and the surfaces $\bar{S}_\alpha$, $\bar{S}_{\alpha+\pi/2}$ are reflections of one another in the point of symmetry of the family. According to (2), this is the only case in which the map of S on $\bar{S}$ reduces to a reflection.

THEOREM 6. *If two surfaces are applicable so that both curvatures are preserved, there exist just two families of curves on each surface which are mapped with preservation of their normal curvatures. The two families form an orthogonal system, namely, the system consisting of the curves which bisect the angles between the lines of curvature associated with a chosen principal normal curvature and the curves corresponding to the lines of curvature on the other surface associated with the same principal normal curvature. The geodesic torsion of each family is equal to that of the family on the other surface to which it does not correspond; in other words, the geodesic torsions of the two families are interchanged by the map.*

If $1/\mathfrak{t} = 1/\bar{\mathfrak{t}}$ and $\alpha \neq 0$, it follows from (10) that $\theta = \alpha/2 - \pi/4$ or $\theta = \alpha/2 + \pi/4 \pmod{\pi}$. Moreover, $1/\mathfrak{r}$ for $\theta = \alpha/2 - \pi/4$ is equal to $1/\bar{\mathfrak{r}}$ for $\theta = \alpha/2 + \pi/4$ and $1/\mathfrak{r}$ for $\theta = \alpha/2 + \pi/4$ is equal to $1/\bar{\mathfrak{r}}$ for $\theta = \alpha/2 - \pi/4$.

THEOREM 7. *If two surfaces are applicable so that both curvatures are preserved, there exist just two families of curves on each surface which are mapped with preservation of their geodesic torsions. The two families form an orthogonal system, namely, the system consisting of the curves which bisect the angles between the two families of curves which are mapped so that their normal curvatures are preserved. The normal curvature of each family is equal to that of the family on the other surface to which it does not correspond; in other words, the normal curvatures of the two families are interchanged by the map.*

The orthogonal systems of Theorems 6 and 7 we shall call respectively the *primary* and *secondary orthogonal systems* of the map of the surfaces on one another. On each surface the curves of either system bisect the angles between the curves of the other system. The map preserves the normal curvatures and interchanges the geodesic torsions of the primary system, and preserves the geodesic torsions and interchanges the normal curvatures of the secondary system.

THEOREM 8. *Two surfaces are applicable with preservation of both curvatures, if and only if they are applicable so that there exist corresponding orthogonal systems whose normal curvatures are preserved and whose geodesic torsions are interchanged, or if and only if they are applicable so that corresponding orthogonal systems exist whose normal curvatures are interchanged and geodesic torsions preserved.*

The conditions have already been proved necessary. That they are sufficient follows immediately from the expressions[16]

$$K = \frac{1}{\mathfrak{r}\mathfrak{r}'} - \frac{1}{\mathfrak{t}^2}, \qquad K' = \frac{1}{\mathfrak{r}} + \frac{1}{\mathfrak{r}'} \tag{11}$$

for the total and mean curvatures of a surface in terms of the normal curvatures $1/\mathfrak{r}$, $1/\mathfrak{r}'$ and geodesic torsions $1/\mathfrak{t}$, $1/\mathfrak{t}' = -1/\mathfrak{t}$ of the two families of curves of an orthogonal system.

Formulas (11) suggest the possibility of corresponding orthogonal systems on S and $\bar{S}$ whose normal curvatures and geodesic torsions are both interchanged. From equations (10) for $\theta = \theta$ and $\theta = \theta + \pi/2$ it is found that only when

[16] See B.M., p. 49; I.M., p. 516.

$\alpha = \pi/2$ do orthogonal systems with this property exist, and that then every pair of corresponding orthogonal systems has the property.

THEOREM 9. *In the case of two surfaces of the same constant mean curvature which are applicable with preservation of the lines of curvature, the normal curvatures and geodesic torsions of each pair of corresponding orthogonal systems are interchanged. This is the only case in which there exist corresponding orthogonal systems on S and $\bar{S}$ whose normal curvatures and geodesic torsions are both interchanged.*

The first part of the theorem seems to deny the existence of the primary and secondary orthogonal systems on the surfaces in question. The paradox is, however, readily explained. The secondary orthogonal systems consist of the lines of curvatures and their torsions are all zero and so are both interchanged and preserved. Similarly, the normal curvatures of the primary orthogonal systems, which consist of the bisectors of the angles between the lines of curvature, are all equal and hence are both preserved and interchanged.

Oblique systems. We now choose a second pair of corresponding families of curves, consisting of the curves $\mathfrak{C}'$ on S, with the slope-angle θ', normal curvature $1/\mathfrak{r}'$, and geodesic torsion $1/\mathfrak{t}'$, and the curves $\bar{\mathfrak{C}}'$ on $\bar{S}$, with slope-angle $\theta' - \alpha$, normal curvature $1/\bar{\mathfrak{r}}'$, and geodesic torsion $1/\bar{\mathfrak{t}}'$, and consider the corresponding systems of curves $\mathfrak{C}$, $\mathfrak{C}'$ and $\bar{\mathfrak{C}}$, $\bar{\mathfrak{C}}'$.

The total and mean curvatures of S and $\bar{S}$ are expressible in terms of the quantities pertaining to either of these two systems of curves. Expressions for them in terms of $1/\mathfrak{r}$, $1/\mathfrak{r}'$, $1/\mathfrak{t}$, $1/\mathfrak{t}'$, and $\omega = \theta' - \theta$ are[17]

$$\text{(12)} \qquad \sin^2 \omega K = \frac{1}{\mathfrak{r}\mathfrak{r}'} - \frac{1}{\mathfrak{p}^2}, \qquad \sin^2 \omega K' = \frac{1}{\mathfrak{r}} + \frac{1}{\mathfrak{r}'} - \frac{2}{\mathfrak{p}} \cos \omega,$$

where $1/\mathfrak{p}$ has either of the values

$$\text{(13)} \qquad \frac{1}{\mathfrak{p}} = \frac{\cos \omega}{\mathfrak{r}} - \frac{\sin \omega}{\mathfrak{t}}, \qquad \frac{1}{\mathfrak{p}} = \frac{\cos \omega}{\mathfrak{r}'} + \frac{\sin \omega}{\mathfrak{t}'},$$

and those in terms of $1/\bar{\mathfrak{r}}$, $1/\bar{\mathfrak{r}}'$, $1/\bar{\mathfrak{p}}'$, and $\bar{\omega} = \omega$ are entirely analogous.

If the two systems of curves are orthogonal ($\omega = \pm \pi/2$), the equal values of $1/\mathfrak{p}$ in (13) reduce to $\mp 1/\mathfrak{t}$ and $\pm 1/\mathfrak{t}'$, and (12) becomes (11). Thus $1/\mathfrak{p}$ assumes the rôle for an arbitrary system which is played by the geodesic torsion for an orthogonal system.

This fact suggests that we seek a generalization of Theorem 7 by demanding that $1/\mathfrak{p} = 1/\bar{\mathfrak{p}}$, $\alpha \neq 0$. From equations (10), (13) and the corresponding equations for $\mathfrak{C}'$, $\bar{\mathfrak{C}}'$, we find that

$$\frac{1}{\mathfrak{p}} = \frac{1}{2}[K' \cos \omega + L \cos (\theta' + \theta)], \qquad \frac{1}{\bar{\mathfrak{p}}} = \frac{1}{2}[K' \cos \omega + L \cos (\theta' + \theta - 2\alpha)].$$

Thus if $1/\mathfrak{p} = 1/\bar{\mathfrak{p}}$ and $\alpha \neq 0$, it follows that $\theta' + \theta = \alpha \pmod{\pi}$ and hence that $\theta = \alpha/2 - \omega/2$, $\theta' = \alpha/2 + \omega/2 \pmod{\pi}$. It is readily shown, then, that $1/\mathfrak{r} = 1/\bar{\mathfrak{r}}'$ and $1/\mathfrak{r}' = 1/\bar{\mathfrak{r}}$.

THEOREM 10. *If two surfaces are applicable so that both curvatures are pre-*

[17] See B.M., pp. 74, 75.

served, there exist an infinity of systems of curves on each surface, depending on an arbitrary function of two variables, every one of which is mapped so that the quantity $1/\mathfrak{p}$ *is preserved. The curves of the primary orthogonal system of the map bisect the angles between those of each of the infinity of systems, so that these systems may be said to form an involution whose double system is the primary orthogonal system and whose orthogonal system is the secondary orthogonal system. The normal curvatures of each of these systems are interchanged by the map.*

The common value of $1/\mathfrak{p}$ and $1/\bar{\mathfrak{p}}$ for the corresponding systems whose curves meet under the angle ω is evidently

$$\frac{1}{\mathfrak{p}} = \frac{1}{\bar{\mathfrak{p}}} = \frac{1}{2}(K' \cos \omega + L \cos \alpha). \tag{14}$$

5. **Conjugate systems and asymptotic lines.**[18] It follows from the opening statement of §4 that the angles between the two families of asymptotic lines on S are equal to those between the two families of asymptotic lines on $\bar{S}$, and that the geodesic torsions of the two families on S are equal to those of the two families on $\bar{S}$. We recall also that the geodesic torsions of the two families on either surface are negatives of one another.

If the asymptotic lines on the two surfaces correspond, they coincide on each surface with the curves of the primary orthogonal systems, by Theorem 6, and hence both surfaces must be minimal. Moreover, since the asymptotic lines correspond, so do also the lines of curvature and therefore $\alpha = \pi/2$. Finally, by Theorem 6, the geodesic torsions of the two families of asymptotic lines on either surface are interchanged by the map and hence are actually preserved except for signs.

The same results follow directly from (14). For the asymptotic lines on the two surfaces correspond if and only if conjugate systems always correspond, and a system of curves on a surface is conjugate if and only if the associated quantity $1/\mathfrak{p}$ vanishes. But, by (14), $1/\mathfrak{p} = 1/\bar{\mathfrak{p}} = 0$ for every value of ω when and only when $K' = 0$ and $\alpha = \pi/2$.

Suppose now that there is just one family of asymptotic lines on each surface which corresponds to a family of asymptotic lines on the other. By Theorem 6, the two families belong, respectively, to the primary orthogonal systems on the two surfaces, and the geodesic torsion of the one is the negative of that of the other. Since the slope angles of the families of the primary orthogonal systems on S are $\alpha/2$ and $\alpha/2 + \pi/2$, it follows that the angles between the asymptotic lines on S, and hence on $\bar{S}$, are α, $\pi - \alpha$.

It is readily shown that if $1/r$ and $1/r'$ are the common normal curvatures of the primary orthogonal system,

$$\frac{4}{rr'} = K'^2 - L^2 \cos^2 \alpha.$$

[18] We exclude developable surfaces in this paragraph.

Hence the case under discussion is characterized by the relation

$$K'^2 - L^2 \cos^2 \alpha = 0, \qquad K' \neq 0.$$

If neither family of asymptotic lines on the one surface corresponds to a similar family on the other, there is a unique conjugate system on each surface which corresponds to a conjugate system on the other. This conjugate system is one of the systems described in Theorem 10, namely, that one, according to (14), whose angle ω is determined by the equation[19]

$$K' \cos \omega + L \cos \alpha = 0.$$

THEOREM 11. *If two surfaces are applicable with preservation of both curvatures, there exists, in general, at least one conjugate system on each surface which corresponds to a conjugate system on the other. Corresponding conjugate systems belong respectively to the involutions on the two surfaces described in Theorem* 10 *and hence their normal curvatures, if existent, are interchanged by the map. In the exceptional case, there exists a single family of asymptotic lines on each surface which corresponds to a family of asymptotic lines on the other. These two families are double families in the two involutions in question and their torsions are negatives of one another.*

Inasmuch as two corresponding systems are conjugate systems if and only if they belong respectively to the involutions on the two surfaces described in Theorem 10, and also to the involutions whose double systems consist of the asymptotic lines, we may also draw the following conclusion.

COROLLARY. *Unique corresponding conjugate systems consist of conjugate-imaginary families of curves if and only if the two surfaces are of negative curvature and the families of asymptotic lines on either separate the constituent families of the primary orthogonal system.*

THEOREM 12. *A necessary and sufficient condition that two non-minimal surfaces be applicable with preservation of both curvatures is that they be applicable either* (a) *so that there exist corresponding conjugate systems which are mapped with interchange of their normal curvatures or* (b) *so that there exist unique corresponding families of asymptotic lines whose geodesic torsions are negatives of one another, provided the smallest positive directed angles from these asymptotic lines to those of the second families are supplementary.*

The necessity of the condition has already been established. The sufficiency follows from (12) in the case of the corresponding conjugate systems and, in the remaining case, from the expressions[20]

$$K = -\frac{1}{t^2}, \qquad K' = -\frac{2}{t} \cot \psi$$

[19] The case of minimal surfaces is excluded. The corresponding conjugate systems consist in this case of the isotropic curves.

[20] See B.M., p. 85.

for the total and mean curvature of a surface in terms of the geodesic torsion of a family of asymptotic lines and the directed angle ψ from these asymptotic lines to those of the second family.

6. **Map referred to fundamental orthogonal systems.** Let the surface $S: x = x(u, v)$ be applicable to the surface $\bar{S}: \bar{x} = \bar{x}(u, v)$ with preservation of both curvatures. Suppose that S is referred to an orthogonal system of directed curves C, C', where the direction of rotation from the directed curve C through an arbitrary point to the directed curve C' is positive, and let the canonical differential equations[21] of the curves C and C' be respectively

$$(15) \qquad Adu + Bdv = 0, \qquad A'du + B'dv = 0.$$

The three fundamental forms of S are[22]

$$(16) \qquad \begin{gathered} ds^2 + ds'^2, \\ \frac{1}{r}\,ds^2 - \frac{2}{\tau}\,dsds' + \frac{1}{r'}\,ds'^2, \\ \left(\frac{1}{r^2} + \frac{1}{\tau^2}\right) ds^2 - \frac{2}{\tau}\,K'dsds' + \left(\frac{1}{r'^2} + \frac{1}{\tau^2}\right) ds'^2, \end{gathered}$$

where $ds = A'du + B'dv, ds' = Adu + Bdv$, and $1/r, 1/r'$ and $1/\tau, 1/\tau'(= -1/\tau)$ are respectively the normal curvatures and geodesic torsions of the curves C, C'; and the Codazzi equations of S are[23]

$$(17) \qquad \begin{gathered} \frac{\partial K'}{\partial s'} + \frac{\partial L}{\partial s'} - \frac{2}{\rho}\,L + \frac{\partial}{\partial s}\left(\frac{2}{\tau}\right) + \frac{2}{\rho'}\,\frac{2}{\tau} = 0, \\ \frac{\partial K'}{\partial s} - \frac{\partial L}{\partial s} - \frac{2}{\rho'}\,L + \frac{\partial}{\partial s'}\left(\frac{2}{\tau}\right) - \frac{2}{\rho}\,\frac{2}{\tau} = 0, \end{gathered}$$

where $\partial/\partial s$, $\partial/\partial s'$ denote directional differentiation in the positive directions of the directed curves C, C' respectively, $1/\rho$, $1/\rho'$ are the geodesic curvatures of these directed curves, and

$$K' = \frac{1}{r} + \frac{1}{r'}, \qquad L = \frac{1}{r} - \frac{1}{r'}.$$

Suppose now that the orthogonal system of curves C, C' is the primary system on S with reference to the map of S on $\bar{S}$. Then if $\bar{K}', \bar{L}, 1/\bar{\tau}$ have the same meanings for the curves $\bar{C}, \bar{C}'$ of the primary orthogonal system on $\bar{S}$ as have $K', L, 1/\tau$ for the curves C, C' on S, we have

$$\bar{K}' = K', \qquad \bar{L} = L, \qquad \frac{1}{\bar{\tau}} = -\frac{1}{\tau},$$

[21] See B.M., p. 48; I.M., p. 501.
[22] See B.M., p. 51; I.M., p. 507.
[23] See B.M., p. 53; I.M., p. 508.

and the Codazzi equations of $\bar{S}$, referred to the curves $\bar{C}$, $\bar{C}'$, become

$$(18)\qquad \begin{aligned} &\frac{\partial K'}{\partial s'} + \frac{\partial L}{\partial s'} - \frac{2}{\rho}L - \frac{\partial}{\partial s}\left(\frac{2}{\tau}\right) - \frac{2}{\rho'}\frac{2}{\tau} = 0, \\ &\frac{\partial K'}{\partial s} - \frac{\partial L}{\partial s} - \frac{2}{\rho'}L - \frac{\partial}{\partial s'}\left(\frac{2}{\tau}\right) + \frac{2}{\rho}\frac{2}{\tau} = 0. \end{aligned}$$

From equations (17) and (18) we find that

$$(19)\qquad \frac{\partial}{\partial s}\left(\frac{1}{\tau}\right) + \frac{2}{\rho'}\frac{1}{\tau} = 0, \qquad \frac{\partial}{\partial s'}\left(\frac{1}{\tau}\right) - \frac{2}{\rho}\frac{1}{\tau} = 0.$$

Conversely, from equations (19), in conjunction with (17), follow equations (18).

Equations (19) constitute conditions necessary and sufficient that $|1/\tau|^{\frac{1}{2}}$ be a common integrating factor[24] of the canonical differential equations (15) of the families of curves[25] C, C'.

THEOREM 13. *There exists a surface applicable to a given surface S with preservation of both curvatures if and only if there are two mutually orthogonal families of curves on S whose canonical differential equations have as a common integrating factor the square root of the absolute value of the geodesic torsion of either of the families.*

COROLLARY. *There are ∞^1 surfaces applicable to S as prescribed when and only when S is an isometric surface containing an orthogonal system of the type described; S then contains ∞^1 such orthogonal systems.*

The corollary follows from the fact that, if an isometric surface admits one surface applicable to it as required, it admits infinitely many, corresponding to ∞^1 values of the parameter α in §2.

We have remarked that equations (19) are equivalent to the relations

$$|1/\tau|^{\frac{1}{2}}(A\,du + B\,dv) = dv_1, \qquad |1/\tau|^{\frac{1}{2}}(A'du + B'dv) = du_1,$$

where u_1, v_1 are functions of u, v. Thus the condition of Theorem 13 is identical with the demand that there exist on S an (isometric) orthogonal system with reference to which the linear element can be written in the form

$$|\tau|\,(du_1^2 + dv_1^2).$$

Suppose now that the curves C, C' on S and the corresponding curves $\bar{C}$, $\bar{C}'$ on $\bar{S}$ constitute the secondary orthogonal systems of the map of S on $\bar{S}$. Then

$$\bar{K}' = K', \qquad \bar{L} = -L, \qquad \frac{1}{\bar{\tau}} = \frac{1}{\tau},$$

[24] The quantity $1/\tau$ is never zero, since the primary orthogonal system never consists of the lines of curvature; see §4.

[25] See B.M., p. 54; I.M., p. 513.

and the Codazzi equations for $\bar{S}$, referred to the curves $\bar{C}$, $\bar{C}'$, become

$$\begin{aligned}(20)\qquad &\frac{\partial K'}{\partial s'} - \frac{\partial L}{\partial s'} + \frac{2}{\rho}L + \frac{\partial}{\partial s}\left(\frac{2}{\tau}\right) + \frac{2}{\rho'}\frac{2}{\tau} = 0,\\ &\frac{\partial K'}{\partial s} + \frac{\partial L}{\partial s} + \frac{2}{\rho'}L + \frac{\partial}{\partial s'}\left(\frac{2}{\tau}\right) - \frac{2}{\rho}\frac{2}{\tau} = 0.\end{aligned}$$

Equations (17) and (20) are obviously equivalent to equations (17) and

$$(21)\qquad \frac{\partial L}{\partial s} + \frac{2}{\rho'}L = 0, \qquad \frac{\partial L}{\partial s'} - \frac{2}{\rho}L = 0.$$

But equations (21) characterize $|L|^{\frac{1}{2}}$ as a common integrating factor[26] of the canonical differential equations (15). Thus we conclude:

THEOREM 14. *There exists a surface applicable to a given surface S with preservation of both curvatures if and only if there are two mutually orthogonal families of curves on S whose canonical differential equations have as a common integrating factor the square root of the absolute value of the difference of the normal curvatures of the two families.*

COROLLARY. *There are ∞^1 surfaces applicable to S as prescribed when and only when S is an isometric surface containing an orthogonal system of the type described; S then contains ∞^1 such orthogonal systems.*

The condition of Theorem 14 is equivalent to demanding that there exist on S an (isometric) orthogonal system with reference to which the linear element can be written in the form

$$\frac{1}{|L|}(du^2 + dv^2).$$

THEOREM 15. *If a surface S admits a surface applicable to it by preservation of both curvatures, the finite equations of the families of curves constituting the primary and secondary orthogonal systems of the map on S can be found by quadratures.*

The theorem follows immediately from the foregoing considerations when one realizes that in both cases the canonical differential equations (15) can be found by algebraic processes.[27]

THEOREM 16. *The spherical representations of the primary orthogonal systems on two surfaces which are applicable with preservation of both curvatures are mapped on one another with preservation of arc lengths.*

This proposition furnishes a characterization of the primary orthogonal systems of the map except when the two surfaces are minimal. It follows from the third of the fundamental forms (16).

By means of the second of these fundamental forms and the relation $K = 1/(rr') - 1/\tau^2$, the following fact is readily proved.

[26] L is never zero, since the secondary orthogonal system never consists of the curves bisecting the angles between the lines of curvature.

[27] See I.M., p. 516.

THEOREM 17. *The asymptotic lines on each surface correspond to a conjugate system on the other if and only if the total curvature of either surface is equal to twice the product of the normal curvatures of the primary orthogonal system.*

7. Surfaces of constant mean curvature.

THEOREM 18. *If there exist on a surface two mutually orthogonal families of curves whose canonical differential equations have both of the common integrating factors* $|1/\tau|^{\frac{1}{2}}$ *and* $|L|^{\frac{1}{2}}$, *the surface has constant mean curvature and the two families cut the lines of curvature under constant angles.*

For, if equations (19) and (21) both hold, $K' = \text{const.}$, by (17), and $1/\tau$ and L have a constant ratio. The directed angle β from the lines of curvature C_1 (see §2) to the curves C is then constant, by virtue of the relation[28]

$$\frac{2}{\tau} = L \tan 2\beta. \tag{22}$$

THEOREM 19. *The canonical differential equations of each two mutually orthogonal families of curves on a surface of constant mean curvature which cut the lines of curvature under constant angles have both* $|1/\tau|^{\frac{1}{2}}$ *and* $|L|^{\frac{1}{2}}$ *as common integrating factors, provided merely that neither quantity vanishes.*[29]

Since $K' = \text{const.}$, the Codazzi equations (17) become

$$\frac{\partial L}{\partial s'} - \frac{2}{\rho}L + \frac{\partial}{\partial s}\left(\frac{2}{\tau}\right) + \frac{2}{\rho'}\frac{2}{\tau} = 0,$$

$$\frac{\partial L}{\partial s} + \frac{2}{\rho'}L - \frac{\partial}{\partial s'}\left(\frac{2}{\tau}\right) + \frac{2}{\rho}\frac{2}{\tau} = 0.$$

From these equations and equation (22), in which β is now constant, equations (19) and (21) are readily deduced and thus the theorem is proved.

In view of the results of §6, Theorems 18, 19 tell us that a surface S admits two surfaces applicable to it with preservation of both curvatures so that a prescribed orthogonal system on it is the primary orthogonal system of the one map and the secondary orthogonal system of the other if and only if S is of constant mean curvature and the given orthogonal system makes constant angles with the lines of curvature.

With reference to an orthogonal system O of the type described we may define four transformations of a surface S of constant mean curvature which carry S into surfaces isometrically mapped on S with preservation of both curvatures, namely, the transformations:

T_1, for which O is the primary orthogonal system on S;
T_2, for which O is the secondary orthogonal system on S;
T_3, which preserves the lines of curvature;
T_0, the identity.

[28] See B.M., p. 54; I.M., p. 516.

[29] For the lines of curvature, $1/\tau = 0$, and for the curves bisecting the angles between the lines of curvature, $L = 0$.

These four transformations, applied to S and O, give rise to four surfaces, S_1, S_2, S_3, $S_0 = S$, and four orthogonal systems, O_1, O_2, O_3, $O_0 = O$, lying respectively on these surfaces.

According to Theorems 6, 7, 9, the four transformations have the following effects on the normal curvatures, $1/r$, $1/r'$, and geodesic torsions, $1/\tau$, $1/\tau'$, of the orthogonal system O:

T_1 preserves $1/r$, $1/r'$ and interchanges $1/\tau$, $1/\tau'$;
T_2 interchanges $1/r$, $1/r'$ and preserves $1/\tau$, $1/\tau'$;
T_3 interchanges $1/r$, $1/r'$ and interchanges $1/\tau$, $1/\tau'$;
T_0 preserves $1/r$, $1/r'$ and preserves $1/\tau$, $1/\tau'$.

It is evident that, with respect to these properties, the four transformations form a group. In particular, if i, j, k are the numbers 1, 2, 3 in any order, $T_i T_j = T_k$ and hence the surface S_i and the orthogonal system O_i are carried into the surface S_j and the orthogonal system O_j by the transformation T_k. We may, then, draw the following conclusions.

THEOREM 20. *The set of four surfaces and associated orthogonal systems S_i, O_i is closed with respect to the group of transformations T_i, $i = 0, 1, 2, 3$. Application of one of the transformations, other than T_0, to the four surfaces and systems interchanges them in pairs, and application of all four transformations to one surface and system yields all four.*

The surfaces S_1, S_2, S_3 belong to the family of surfaces $\bar{S}$ which are applicable to $S = S_0$ with preservation of both curvatures. If, as in §3, we denote the surface $\bar{S}$ for which $\alpha = \alpha$ by $\bar{S}_\alpha$, $0 \leqq \alpha < \pi$, and denote by θ, $0 \leqq \theta < \pi/2$, the smallest non-negative directed angle by which the orthogonal system O on S_0 is advanced over the lines of curvature, then S_0, S_1, S_2, S_3 are respectively the surfaces S, $\bar{S}_{2\theta}$, $\bar{S}_{2\theta\pm\pi/2}$, $\bar{S}_{\pi/2}$. As θ varies from 0 to $\pi/2$, the two variable surfaces of the set each trace the entire family of surfaces $\bar{S}_\alpha$ just once.

As the comment on Theorem 9 suggests, there are two special positions of O for which the surfaces S_i, and the transformations T_i, coincide in pairs. When O consists of the lines of curvature on S_0, then $\theta = 0$, so that $S_1 = S_0$, $S_2 = S_3$ and $T_1 = T_0$, $T_2 = T_3$; and when O consists of the bisectors of the angles between the lines of curvature, then $\theta = \pi/4$, $S_2 = S_0$, $S_1 = S_3$ and $T_2 = T_0$, $T_1 = T_3$.

HARVARD UNIVERSITY.

THE IDEAL WARING THEOREM FOR TWELFTH POWERS

By L. E. Dickson

1. **Ideal.** Let g denote the greatest integer $< (3/2)^n$. Let

$$(1) \qquad P = g2^n - 1 , \qquad I = g + 2^n - 2 .$$

Evidently $P < 3^n$. Consider all the ways in which P can be a sum of positive integral n-th powers, necessarily 1^n or 2^n. There will occur more than I summands except when there are exactly $g - 1$ terms 2^n and exactly $2^n - 1$ terms 1. Hence P is a sum of I, but not fewer, n-th powers.

When $n = 2, 3, 4$, $g = 2, 3, 5$, $I = 4, 9, 19$. Lagrange and Euler proved that every positive integer is a sum of four squares. In 1770 Waring conjectured that "every positive integer is a sum of 9 cubes, also of 19 fourth powers, etc." The proof for 9 cubes was first obtained by Wieferich.[1] For fourth powers, the best result yet proved is that 35 suffice, while all integers $< 10^{26}$ are sums of 19. All tables extant confirm the following amplification of Waring's conjecture: *Every positive integer is a sum of $I = g + 2^n - 2$ integral n-th powers* $\geqq 0$. This will be called the ideal Waring theorem. Proof has been published only for $n = 2$ and $n = 3$.

2. **Summary for twelfth powers.** We here prove the ideal Waring theorem for $n = 12$, viz.,

Theorem 1. *Every positive integer is a sum of $I = 4223$ integral twelfth powers.*

But we go further and prove

Theorem 2. *Every integer $> 2 \cdot 3^{12}$ is a sum of 2405 twelfth powers. Every one $> 3 \cdot 3^{12}$ is a sum of 1560 powers. Every one $> 2 \cdot 5^{12} + 7^{12} + 8^{12}$ is a sum of 440.*

There are only two earlier results.[2] From Kempner's identity, we find that 6 1/4 billion twelfth powers suffice. By a short table, the writer proved also that 10711 suffice. The present table gives decompositions of 6 908 733 consecutive integers into twelfth powers.

3. **Minimum decompositions.** We employ

$$a = 2^{12} = 4096 , \qquad b = 3^{12} = 531\,441 , \qquad c = 4^{12} = 16\,777\,216 ,$$

$$(1) \quad d = 5^{12} = 244\,140\,625 , \qquad f = 6^{12} = 2\,176\,782\,336 ,$$

$$g = 7^{12} = 13\,841\,287\,201 , \qquad h = 8^{12} = 68\,719\,476\,736 ,$$

Received February 7, 1936.

[1] Math. Annalen, vol. 66 (1909), pp. 99–101. The proof was corrected and greatly simplified by Dickson, Trans. Amer. Math. Soc., vol. 30 (1927), pp. 1–7.

[2] Dickson, Bull. Amer. Math. Soc., vol. 39 (1933), p. 713.

$$(2) \qquad b = 129a + 3057\,, \qquad c = 31b + 73a + 3537\,,$$

$$(3) \qquad d = 14c + 17b + 55a - 176\,,$$

$$(4) \qquad f = 8d + 13c + 10b + 58a + 1550\,,$$

$$(5) \qquad g = 6f + 3d + 2c + 27b + 65a + 1731\,,$$

$$(6) \qquad h = 4g + 6f + d + 2c + 30b - a - 275\,.$$

By a *resolution* of n we mean one of the ways of expressing n as a linear function of $a, b, c, \cdots$, whose coefficients are integers. In case all the coefficients are $\geqq 0$, the resolution is a *decomposition*, and their sum is called the *weight* w. Thus, (2), (4), (5) are decompositions, while (3) and (6) are not, but are resolutions. A decomposition of n is *minimal* if there is no other decomposition of n of smaller weight.

For $s = 2d + g + h$, we seek a minimum decomposition of every integer between s and $s + d$. Such an integer can be expressed in the form $r + N$, where $0 \leqq r < a$, and

$$(7) \quad N = Aa + Bb + Cc + s\,, \quad 0 \leqq A \leqq 129\,, \quad 0 \leqq B \leqq 31\,, \quad 0 \leqq C \leqq 14\,.$$

Assign any integral values to A, B, C within their limits. All homogeneous resolutions, involving[3] $g + h$ but not d, of $r + N$ ($r = 0, \cdots, 4095$) are evidently obtained by adding N to all resolutions of $0, \cdots, 4095$ of the form $L - 2d$, where L is linear and homogeneous in a, b, c. One of the latter resolutions is obtained from the double of (3),

$$(8) \qquad 110a + 34b + 28c - 2d = 352\,.$$

Evidently all such resolutions of $0, \cdots, 4095$ are found by adding to (8) all homogeneous resolutions into a, b, c of $-352, \cdots, 3743$. The latter are obviously found by repeated additions of

$$130a - b = 1039\,, \qquad 74a + 31b - c = 559\,,$$

and (3) with terms similarly transposed, together with the negatives of these four.

Actually we desire homogeneous decompositions and not resolutions. Hence in what precedes we retain only those resolutions of $0, \cdots, 4095$ which involve no one of $-130a$, $-32b$, $-15c$.

4. **Leaders.** A decomposition is not minimal if it is the sum of the left member of one of the following equations and any function whose coefficients are all $\geqq 0$.

[3] For those involving $d + g + h$, but not s, use $L - d$. For those involving $2d + g$, but not h, use $L - h$; etc.

$$63b + 18a = 79 + 2c, \qquad 172b + d = 61 + a + 20c,$$

$$31d + 68c + 11a = 11 + 4b + 4f, \qquad 58d + 31a = 25 + 32b + 18c + g,$$

$$2f + 17c + 104a = 12 + b + 19d, \qquad 3f + 28d + 28c + 10b = 37 + 18a + g,$$

$$5f + 71d + 3b = 8 + 11a + 32c + 2g, \qquad 7f + 19c + 99a = 92 + 12b + 7d + g,$$

$$2g + 55d + 15c + 21a = 3 + 6b + 19f,$$

$$4g + 79d + 36c = 55 + 9a + 12b + 3f + h,$$

$$h + 34b + 90a = 44 + 9c + 37d + g,$$

$$h + 56d + 24b + 4a = 14 + 18c + 25f + 2g,$$

$$h + 23c + 111b = 39 + a + 54d + 13f + 2g.$$

By convention a like result holds for the following equations whose two members are of equal weight:

$$130c = 120 + a + 8b + f, \qquad 15d + 40c + 38b + 42a = 133 + 2f,$$

$$15d + 42c + 24a = 54 + 25b + 2f, \qquad 45d + 73a = 42 + 67b + 4c + 5f,$$

$$2f + 88b = 25 + 6a + 44c + 15d, \qquad 3f + 46c + 42b = 12 + 49a + 30d,$$

$$4f + 117a = 55 + 22b + 9c + 35d, \qquad 6f + 3b + 93a = 1 + 51c + 50d,$$

$$10f + 85c + a = b + 95d.$$

5. **List of equations.** We employ equations $P = r$ $(0 \leqq r \leqq 4095)$ in which the coefficients of f and c are $\geqq 0$, those of g and h are $\geqq -1$, while the coefficient of d is $\geqq -2$ and that of b is $\geqq -3$. In fact, we shall use (7) only when $C = 0$, $B \leqq 3$, and we then require that all coefficients of $N + P$ be $\geqq 0$, whence $N + P$ is a decomposition of r. Since we retain only minimal decompositions, we have discarded $N + P$ if it be the sum of a leader and a function whose coefficients are all $\geqq 0$. Further, $N + P$ were found to be not minimal (and discarded) during the construction of the table in §6. For reasons explained in §10 we have abridged a longer list of such equations. In the list, w is the weight of P in $P = r$. When the coefficients of f, d, c are omitted, they are the same as in the line above.

$-h-g$ plus

No.	f	d	c	b	a	r	w
1	31	61	10	38	−8	754	130
2	32	51	27	−1	26	2577	133
3	32	52	12	13	45	2960	152
4				14	−85	1921	23
5	33	43	13	21	−88	3647	20
6	34	34	13	59	−18	1836	120
7	34	35	0	10	−17	2140	60
8	35	23	45	6	−3	3276	104
9	35	25	15	34	35	4042	142
10				35	−95	3003	13
11	35	25	16	3	−39	3483	38
12	35	26	1	17	−20	3866	57
13	36	15	31	27	12	1289	119
14	36	16	16	41	31	1672	138
15	36	16	17	10	−43	1113	34
16	36	17	2	24	−24	1496	53
17	37	6	31	66	−47	2535	91
18	37	6	33	3	−65	2456	12
19	37	7	16	80	−28	2918	110
20	37	7	17	48	28	3398	135
21				49	−102	2359	6
22	37	7	18	17	−46	2839	31
23	37	8	2	63	−83	2742	25
24	37	8	3	31	−27	3222	50
25	37	8	4	0	−101	2663	−54
26	38	−2	19	24	−50	469	27
27	38	−1	4	38	−31	852	46
28	38	−1	5	6	25	1332	71
29				7	−105	293	−58

$-h+0g$ plus

No.	f	d	c	b	a	r	w
30	25	58	6	74	−55	2564	107
31	26	48	24	3	34	771	134
32	27	39	24	42	−25	2017	106
33	27	39	25	10	31	2497	131
34	27	40	10	25	−80	1841	21
35	28	30	25	49	−28	3743	103
36	28	31	11	32	−83	3567	18
37	28	31	12	0	−27	4047	43
38	29	21	27	24	24	1853	124
39				25	−106	814	−5
40	29	22	11	70	−13	1756	118
41	30	11	41	81	−110	2236	52
42	30	11	42	49	−54	2716	77
43	30	11	43	17	2	3196	102
44	30	12	28	31	21	3579	121
45				32	−109	2540	−8

$-h+0g$ plus (*continued*)

No.	f	d	c	b	a	r	w
46	30	12	29	0	−53	3020	17
47	31	2	43	56	−58	346	73
48	31	2	44	24	−2	826	98
49	31	3	28	70	−39	729	92
50	31	3	29	38	17	1209	117
51				39	−113	170	−12
52	31	3	30	6	73	1689	142
53				7	−57	650	13
54	31	4	14	53	−94	553	7
55	31	4	15	21	−38	1033	32
56	31	5	0	35	−19	1416	51
57	31	5	1	3	37	1896	76
58				4	−93	857	−53

$-h+g$ plus

No.	f	d	c	b	a	r	w
59	18	63	18	25	−6	3225	118
60	20	45	20	39	−13	2581	111
61	20	46	6	22	−68	2405	26
62	21	33	66	4	−73	3158	51
63	21	36	21	46	−17	211	107
64	21	37	6	60	2	594	126
65	22	28	7	67	−1	2320	123
66	22	28	8	36	−75	1761	19
67	23	14	83	4	−99	2131	25
68	23	16	54	0	−5	3377	88
69	23	17	37	78	−98	2800	57
70	23	17	39	14	14	3760	107
71	23	18	23	60	−23	3663	101
72	23	19	10	11	−22	3967	41
73	24	8	40	22	−120	351	−26
74	24	9	24	67	−27	1293	97
75	24	9	25	35	29	1773	122
76	24	9	26	4	−45	1214	18
77	24	10	9	81	−8	1676	116
78	24	10	10	50	−82	1117	12
79	24	10	11	18	−26	1597	37
80	25	−1	41	28	7	3116	100
81	25	0	25	74	−30	3019	94
82	25	0	27	11	−48	2940	15
83	25	1	12	25	−29	3323	34

$-h+2g$ plus

No.	f	d	c	b	a	r	w
84	10	77	28	1	−10	1680	107
85	12	58	45	1	−36	653	81
86	12	60	15	29	2	1419	119

$-h+2g$ plus (*continued*)

No.	f	d	c	b	a	r	w
87	12	61	1	12	−53	1243	34
88	13	49	46	8	−39	2379	78
89	13	51	16	36	−1	3145	116
90	13	51	17	5	−75	2586	12
91	13	52	1	50	18	3528	135
92	13	52	2	19	−56	2969	31
93	14	43	2	57	14	1158	131
94	14	43	3	26	−60	599	27
95	15	32	34	4	29	2598	115
96	15	33	19	18	48	2981	134
97				19	−82	1942	5
98	15	34	4	33	−63	2325	24
99	15	34	5	1	−7	2805	49
100	16	20	79	1	−87	2695	30
101	16	23	34	43	−30	3844	87
102	16	23	35	11	25	228	111
103	16	24	19	57	−12	131	105
104	16	25	4	71	7	514	124
105	16	25	6	8	−11	435	45
106	17	13	51	4	3	1571	89
107	17	15	22	1	−33	1778	23
108	17	16	5	78	4	2240	121
109	17	16	7	15	−14	2161	42
110	18	1	96	1	−113	1668	4
111	18	3	67	−3	−19	2914	67
112	18	4	52	11	0	3297	86
113	18	5	37	25	19	3680	105
114				26	−111	2641	−24
115	18	6	21	71	−18	3583	99
116	18	6	23	8	−36	3504	20
117	18	7	6	85	1	3966	118
118	18	7	7	54	−73	3407	14
119	19	−2	8	61	−77	1037	10
120	19	−2	9	29	−21	1517	35

$-h+3g$ plus

No.	f	d	c	b	a	r	w
121	2	93	8	5	24	901	134
122	4	74	25	5	−1	3970	109
123	5	65	26	12	−5	1600	105
124	6	55	42	5	−27	2943	83
125	7	48	14	8	63	1819	142
126				9	−67	780	13
127	8	36	59	5	−53	1916	57
128	8	39	15	15	60	3545	139
129				16	−70	2506	10
130	8	40	0	30	−51	2889	29
131	9	29	31	8	37	792	116

$-h+3g$ plus (*continued*)

No.	f	d	c	b	a	r	w
132				9	−92	3849	−12
133	9	30	15	54	0	695	110
134	9	30	16	23	−74	136	6
135	9	31	1	37	−55	519	25
136	9	31	2	5	1	999	50
137	10	17	75	36	−5	1448	135
138	10	17	76	5	−79	889	31
139	10	18	61	19	−60	1272	50
140	10	19	45	65	−97	1175	44
141	10	19	47	1	15	2135	94
142	10	20	31	47	−22	2038	88
143	10	20	32	15	34	2518	113
144				16	−96	1479	−16
145	10	21	17	29	53	2901	132
146				30	−77	1862	3
147	10	22	2	44	−58	2245	22
148	10	22	3	12	−2	2725	47
149	11	8	77	12	−82	2615	28
150	11	9	62	26	−63	2998	47
151	11	11	33	23	−99	3205	−19
152	11	12	18	37	−80	3588	0
153	11	12	19	5	−24	4068	25
154	11	13	3	51	−61	3971	19
155	11	13	4	19	−6	355	43
156	12	−1	78	19	−86	245	24
157	12	0	64	1	−11	1108	68
158	12	1	49	15	8	1491	87
159				16	−122	452	−42
160	12	3	19	44	−84	1218	−4
161	12	3	20	12	−28	1698	21
162	12	4	4	58	−65	1601	15
163	12	4	5	26	−9	2081	40

$-h+4g$ plus

No.	f	d	c	b	a	r	w
164	1	40	84	5	−5	2273	128
165	1	42	53	65	−23	2559	141
166	1	42	55	2	−41	2480	62
167	1	43	38	79	−4	2942	160
168	1	43	39	48	−78	2383	56
169	1	43	40	16	−22	2863	81
170	1	45	10	44	16	3629	119
171	1	45	11	13	−58	3070	15
172	2	31	85	12	−8	3999	125
173	2	33	56	9	−45	110	58
174	2	34	41	23	−26	493	77
175	2	35	26	37	−7	876	96
176	2	35	27	6	−81	317	−8

$-h+4g$ plus (*continued*)

No.	f	d	c	b	a	r	w
177	2	36	11	51	12	1259	115
178	2	36	12	19	68	1739	140
179				20	−62	700	11
180	3	21	101	5	−31	1246	102
181	3	22	86	19	−12	1629	121
182	3	23	72	2	−67	1453	36
183	3	24	57	16	−48	1836	55
184	3	25	41	62	−85	1739	49
185	3	26	27	44	−10	2602	93
185′	3	27	12	58	9	2985	112
186	3	27	13	27	−65	2426	8
187	4	14	72	40	4	3738	137
188	4	14	73	9	−70	3179	33
189	4	16	42	69	−88	3465	46
190	4	16	43	37	−32	3945	71
191	4	16	44	5	23	329	95
192				6	−106	3386	−33
193	4	17	28	51	−14	232	89
194	4	17	29	19	42	712	114
195				20	−87	3769	−14
196	4	18	14	34	−69	56	4
197	4	18	15	2	−13	536	29
198	4	19	0	16	6	919	48
199	5	4	89	2	−93	426	10
200	5	5	74	16	−74	809	29
201	5	6	59	30	−55	1192	48
202	5	7	44	44	−36	1575	67
203	5	7	45	12	20	2055	92
204				13	−110	1016	−37
205	5	8	29	58	−17	1958	86
206	5	8	30	26	39	2438	111
207				27	−91	1399	−18
208	5	9	15	41	−72	1782	1
209	5	9	16	9	−16	2262	26
210	5	10	0	55	−53	2165	20
211	5	10	1	23	3	2645	45
212	6	−2	46	20	−113	2742	−40
213	6	−1	30	65	−20	3684	83
214	6	−1	31	34	−94	3125	−21
215	6	−1	32	1	91	548	132
216				2	−38	3605	4
217	6	0	15	79	−1	4067	102
218	6	0	16	48	−75	3508	−2
219	6	0	17	16	−19	3988	23
220	6	1	1	62	−56	3891	17
221	6	1	2	30	−1	275	41
222	6	1	3	−2	55	755	66

$-h+5g$ plus

No.	f	d	c	b	a	r	w
223	0	−2	0	2	64	3045	68
224				3	−66	2006	−61

$-g$ plus

No.	f	d	c	b	a	r	w
225	0	54	39	7	−82	2188	17
226	0	55	24	21	−63	2571	36
227	0	56	9	35	−44	2954	55
228	0	56	10	3	12	3434	80
229	1	43	69	17	−49	3707	80
230	1	44	55	0	−104	3531	−5
231	1	45	40	13	44	857	142
232				14	−85	3914	14
233	1	46	25	28	−67	201	32
234	1	47	9	74	−104	104	26
235	1	47	11	10	8	1064	76
236	2	33	85	10	−72	954	57
237	2	35	56	6	22	2200	120
238				7	−108	1161	−9
239	2	36	41	20	41	2583	139
240				21	−89	1544	10
241	2	37	27	3	−14	2407	54
242	2	38	11	49	−51	2310	48
243	2	38	12	17	5	2790	73
244				18	−125	1751	−56
245	3	23	101	3	−94	2297	35
246	3	24	86	17	−75	2680	54
247	3	29	13	24	1	420	69
248	4	14	102	10	−97	4023	32
249	4	16	73	6	−4	1173	94
250	4	18	44	3	−40	1380	28
251	4	19	28	49	−77	1283	22
252	4	19	29	17	−21	1763	47
253	4	20	15	0	−76	1587	−38
254	4	21	0	13	73	3009	110
255				14	−57	1970	−19
256	5	4	118	2	10	2309	138
257	5	6	89	−1	−26	2516	72
258	5	8	59	27	12	3282	110
259	5	9	43	73	−25	3185	104
260	5	9	45	10	−43	3106	25
261	5	10	30	24	−24	3489	44
262	5	11	14	70	−61	3392	38
263	5	11	16	6	50	256	87
264				7	−79	3313	−41
265	5	12	0	52	13	159	81

− g plus (*continued*)								0 plus						
No.	f	d	c	b	a	r	w	No.	d	c	b	a	r	w
266	5	12	1	20	69	639	106	276	−2	26	98	−1	3488	121
267				21	−60	3696	−22	277	−2	27	67	−75	2929	17
268	6	−1	61	3	−66	353	2	278	−2	28	35	−19	3409	42
269	6	0	46	17	−47	736	21	279	−2	29	3	37	3889	67
270	6	1	32	0	−102	560	−64	280			4	−93	2850	−62
271	6	2	17	13	47	1982	84	281	−1	13	49	0	3792	61
272				14	−83	943	−45	282	−1	14	17	55	176	85
273	6	3	1	59	10	1885	78	283			18	−74	3233	−43
274	6	3	2	27	66	2365	103							
275				28	−64	1326	−26							

6. **The table.** Except for the final three tablettes marked $B = 3$, all the tablettes have $B = 0$ and lead to decompositions of $r + Aa + s$, where $s = 2d + g + h$.

Tablette $A = 0$ is merely an arrangement according to the constant term r (second column) of those equations (number in first column) in which the coefficients of a and b are $\geqq 0$. The third column gives $w + D - 1$, where w is the weight of the left member of the equation and D is the difference between its r and the following r. For example,

No.	r	D	w	$w+D-1$
102	228	28	111	138
263	256	73	87	159
191	329		95	

Hence 138 is the weight for the largest (unprinted) r between 228 and 256. For tablette $A = 0$ the largest entry in the third column is the final entry $m = 195$. It was computed from the (unprinted) final $r = a = 4096$ and the weight 142 of No. 9. In (7), N is here s, whose weight is 4. Hence each of the integers $r + s$ $(r = 0, \cdots, 4095)$ is a sum of $195 + 4$ twelfth powers.

Tablette $A = 47$ employs only equations in which the coefficient of b is $\geqq 0$ and that of a is $\geqq -47$.

An earlier table $A = t$ is denoted by (t). A row of dots means insert from the next earlier tablette.

For $A = 8$ we cite only the numbers of the new equations to be inserted in tablettes $A = 0, 1$. The effect of the insertions is usually shown in detail under $A = 19$. Similarly for $A = 10, 43, 53, 60, 63$.

$A = 0$

	0	156
265	159	149
102	228	138
263	256	159
191	329	185
247	420	162
104	514	157
215	548	177
64	594	170
266	639	161
133	695	126
194	712	172
31	771	154
131	792	180
231	857	185
121	901	151
198	919	127
136	999	114
235	1064	169
93	1158	181
50	1209	166
177	1259	144
13	1289	161
28	1332	157
86	1419	190
158	1491	166
106	1571	189
14	1672	154
52	1689	191
178	1739	173
75	1773	167
125	1819	175
38	1853	155
273	1885	88
57	1896	161
271	1982	156
203	2055	171
141	2135	158
237	2200	159
108	2240	189
256	2309	193
274	2365	175
206	2438	169
33	2497	151
143	2518	177
239	2583	153
95	2598	161
211	2645	189
243	2790	183
145	2901	190
3	2960	172
96	2981	161
254	3009	145
223	3045	138
80	3116	179
43	3196	187
258	3282	124
112	3297	186
20	3398	170
228	3434	173
91	3528	151
128	3545	172
44	3579	170
170	3629	169
113	3680	162
187	3738	158
70	3760	138
281	3792	157
279	3889	143
117	3966	193
9	4042	195

$A = 1$

256	2309	148
65	2320	167
274	2365	175
....	....	....
117	3966	121
122	3970	180
9	4042	166
217	4067	130

$A = 8$

155, 48, 175,
249, 137, 123,
77, 164, 148,
99, 167, 89, 59,
8, 68, 172

$A = 10{:}84$

$A = 19$

(0)		0-256
191	329	120
155	355	107
247	420	162
104	514	145
197	536	131
(0)		639-771
131	792	149
48	826	147
175	876	139
198	919	127
136	999	114
235	1064	169
93	1158	145
249	1173	129
50	1209	166
177	1259	144
13	1289	161
28	1332	154
56	1416	125
158	1491	166
106	1571	117
123	1600	133
181	1629	163
14	1672	141
77	1676	119
84	1680	165
178	1739	156
40	1756	134
75	1773	167
125	1819	158
6	1836	136
(0)		1853-2200
108	2240	153
164	2273	163
256	2309	148
65	2320	167
274	2365	144
241	2407	143
33	2497	151
143	2518	175
60	2581	127
95	2598	118
185	2602	135
211	2645	124
148	2725	111
243	2790	87
99	2805	144
145	2901	172
167	2942	177
(0)		2960-3045
80	3116	128
89	3145	166
43	3196	130
59	3225	168
8	3276	124
112	3297	165
68	3377	144
228	3434	133
276	3488	160
(0)		3528-3889
117	3966	121
122	3970	137
172	3999	167
9	4042	166
217	4967	130

$A = 27$

194	712	155
1	754	146
31	771	154
...	...	...
235	1064	119
157	1108	132
249	1173	129
...	...	...
40	1756	124
252	1763	136
(0)	1853-1982	
203	2055	117
163	2081	119
109	2161	142
209	2262	128
274	2365	144
241	2407	143
33	2497	151
143	2518	153
165	2559	162
60	2581	127
...	...	...
99	2805	106
169	2863	160
124	2943	148
254	3009	145
223	3045	138
80	3116	128
89	3145	166
43	3196	127
24	3222	124
112	3297	165
68	3377	119
278	3409	121
261	3489	133
44	3579	124
115	3583	144
170	3629	152
71	3663	117
(0)	3680-3889	
117	3966	118
72	3967	140
217	4067	130

$A = 32$

50	1209	143
180	1246	144
13	1289	122
74	1293	135
28	1332	154
56	1416	125
158	1491	91
16	1496	73
120	1517	114
79	1597	119
84	1680	124
161	1698	85
(27)	1763-3116	
89	3145	155
259	3185	114
43	3196	127
24	3222	124
112	3297	111
83	3323	119
(27)	3409-4067	

$A = 36$

103	131	132
265	159	149
102	228	138
263	256	105
221	275	120
155	355	107
247	420	83
105	435	145
197	536	131
266	639	119
85	653	122
133	695	126
...	...	...
57	1896	137
205	1958	109
271	1982	118
32	2017	143
203	2055	117
...	...	...
169	2863	135
19	2918	134
124	2943	148
...	...	...
113	3680	108
213	3684	141
35	3743	119
70	3760	138
281	3792	134
12	3866	79
279	3889	143

$A = 43$

250, 107, 260

$A = 47$

173	110	106
265	159	97
282	176	119
63	211	123
102	228	114
193	232	112
263	256	105
221	275	120
155	355	107
247	420	83
105	435	102
174	493	119
(36)	536-695	
194	712	130
49	729	98
269	736	110
48	826	123
27	852	112
198	919	127
136	999	114
235	1064	119
157	1108	72
15	1113	129
50	1209	121
76	1214	135
28	1332	118
250	1380	63
(32)	1416-1698	
252	1763	61
107	1778	129
273	1885	88
57	1896	137
205	1958	109
271	1982	118
32	2017	126
142	2038	104
(27)	2055-2262	
274	2365	116
88	2379	105
241	2407	126
166	2480	116
17	2535	136
(19)	2581-2805	
99	2805	82
22	2839	134
124	2943	93
227	2954	119
81	3019	119
223	3045	128
260	3106	114
(32)	3196-3409	
261	3489	58
116	3504	120
216	3605	141
35	3743	119

70	3760	138
281	3792	112
101	3844	108
12	3866	79
279	3889	122
190	3945	92
72	3967	120
37	4047	91

A = 53:210, 229

A = 60:127, 30, 267

$A = 63$

197	536	91
94	599	77
53	650	98
269	736	110
.....		
15	1113	112
201	1192	69
76	1214	46
87	1243	122
28	1332	118
.....		
107	1778	80
183	1836	103
273	1885	88
57	1896	95
127	1916	98
205	1958	109
.....		
109	2161	45
210	2165	116
209	2262	73
242	2310	116
88	2379	105
.....		
17	2535	119
30	2564	113
226	2571	109
211	2645	115
42	2716	85
148	2725	111
243	2790	87
99	2805	82
22	2839	80
130	2889	93
227	2954	119
81	3019	94
46	3020	102
260	3106	114
.....		

216	3605	94
267	3696	73
281	3792	112

$A = 72$

198	919	82
236	954	101
136	999	83
55	1033	106
157	1108	72
.....		
56	1416	87
182	1453	78
16	1496	73
120	1517	114
79	1597	40
162	1601	111
161	1698	85
.....		
271	1982	107
224	2006	93
109	2161	45
.....		
88	2379	103
61	2405	100
.....		
260	3106	97
188	3179	75
24	3222	124

$A = 75$

173	110	83
134	136	70
233	201	105
221	275	111
47	346	81
155	355	107
247	420	83
105	435	78
26	469	76
135	519	41
.....		
269	736	93
200	809	71
27	852	112
.....		
87	1243	62
139	1272	103
275	1326	63
(72)	1416-1496	
120	1517	92
202	1575	88
(72)	1597-2161	

210	2165	99
147	2245	38
209	2262	73
242	2310	62
98	2325	103
61	2405	46
186	2426	37
18	2456	61
129	2506	74
226	2571	50
90	2586	70
211	2645	79
246	2680	89
(63)	2716-2839	
130	2889	68
277	2929	106
81	3019	94
46	3020	102
260	3106	97
188	3179	75
24	3222	60
283	3233	46
83	3323	102
262	3392	52
118	3407	95
261	3489	62
218	3508	94
(63)	3605-3866	
279	3889	68
220	3891	92
72	3967	44
154	3971	94
37	4047	91

$A = 87$

196	56	83
134	136	70
233	201	75
156	245	53
221	275	82
176	317	27
268	353	83
(75)	435-599	
53	650	62
179	700	46
269	736	64
126	780	41
200	809	71
27	852	82
138	889	60
198	919	71
272	943	44
55	1033	35

119	1037	80
157	1108	72
15	1113	37
78	1117	86
201	1192	69
76	1214	21
160	1218	60
251	1283	64
275	1326	63
(75)	1416-1517	
202	1575	78
253	1587	72
161	1698	83
66	1761	35
107	1778	26
208	1782	80
146	1862	82
97	1942	32
255	1970	16
224	2006	72
7	2140	80
109	2161	45
210	2165	42
225	2188	73
(75)	2245-2310	
98	2325	81
168	2383	77
(75)	2405-2645	
246	2680	68
100	2695	59
148	2725	63
23	2742	87
99	2805	82
22	2839	80
130	2889	68
277	2929	27
82	2940	43
92	2969	59
150	2998	68
46	3020	66
171	3070	50
260	3106	76
62	3158	71
188	3179	75
24	3222	60
283	3233	36
264	3313	52
(75)	3407-3605	
267	3696	50
195	3769	82
12	3866	79
(75)	3889-4047	

$A = 94$

197	536	45
54	553	52
94	599	77
. . . . . .		
120	1517	61
240	1544	52
253	1587	72
. . . . . .		
118	3407	71
189	3465	69
261	3489	62
218	3508	77
152	3588	16
216	3605	45
5	3647	68
(87)	3696-3889	
220	3891	39
232	3914	66
72	3967	44
154	3971	35
219	3988	81
37	4047	63
153	4068	52

$A = 98$

78	1117	69
140	1175	60
201	1192	69
. . . . . .		
275	1326	46
207	1399	61
144	1479	48
(94)	1544-2725	
23	2742	82
69	2800	61
(87)	2805-3313	

$A = 113$

196	56	51
234	104	57
134	136	39
51	170	62
156	245	53
221	275	58
29	293	74
199	426	52
26	469	76
135	519	41
197	536	45
54	553	13
270	560	25
(87)	650-780	
200	809	33
39	814	37
58	857	32
272	943	27
204	1016	59
15	1113	37
78	1117	55
238	1161	43
(87)	1214-1283	
(98)	1326-1544	
253	1587	42
110	1668	33
161	1698	61
184	1739	70
66	1761	35
107	1778	26
208	1782	59
34	1841	41
146	1862	61
4	1921	43
97	1942	32
255	1970	16
224	2006	63
67	2131	54
(87)	2161-2310	
98	2325	57
21	2359	51
(87)	2405-2571	
90	2586	40
149	2615	53
114	2641	−3
25	2663	24
212	2742	56
22	2839	41
280	2850	27
(87)	2940-3233	
264	3313	31
192	3386	63
11	3483	62
218	3508	20
230	3531	30
36	3567	38
(94)	3588-3696	
195	3769	65
132	3849	29
(94)	3891-3971	
219	3988	57
248	4023	55
37	4047	63
153	4068	52

$A = 122$

29	293	−1

73	351	48	214	3125	58	145	2901	144
199	426	35	151	3205	8	111	2914	137
159	452	41		3233-4096		185′	2985	137
	536-2165							
225	2188	64		$A = 125$			$B = 3, \quad A = 36$	
41	2236	60	184	1739	60	33	2497	149
147	2245	38	244	1751	33	257	2516	136
209	2262	60	34	1841	41	60	2581	127
245	2297	62						
	2325-2456			$B = 3, \quad A = 19$			$B = 3, \quad A = 43$	
129	2506	43	143	2518	171	194	712	130
45	2540	37	2	2577	136	49	729	117
	2586-3070		60	2581	127	222	755	136
260	3106	43		....		48	826	147

7. **Conclusion from the table.** In the tablette with a fixed A, let m denote the maximum entry of its third column. For $B = 0$ we have[4]

A	0	1-7	8-9	10-26	27-31	32-46	47-52	53-59	60-62
m	195	191	178*	177*	166	162*	142	138	135
$A + m$	195	198	187	203	197	208	194	197	197

A	63-71	72-4	75-86	87-93	94-7	98-112	113-21	122-4	125-9
m	127	124	112	95	87*	83	76*	70*	68
$A + m$	198	198	198	188	184	195	197	194	197

Here $A + m$ is found from the largest A in its sequence. Thus $A + m \leqq 208$ for every A. We have also the further needed facts. For $A = 19$, max $= 172$ except for gaps 175 at 2580 and 177 at 2959. For $A = 36$, max $= 156$ except for gap 162 at 2580. For $A = 43$, max $= 152$ except for preceding gap and for $r = 712$-771.

Employing also tablettes with $B = 3$, we see that, when $B = 3$, $A + m \leqq 198$ for every A. Hence all our results imply that, if $B \leqq 12$, $B + A + m \leqq 210$ for every $A \leqq 129$. Applying (2_1) we obtain, as in §6,

THEOREM 3. *Every integer between $s = 2d + g + h$ and $s + 13b$ is a sum of 214 twelfth powers.*

8. **Ascent.** We take $n = 12$ in the writer's[5]

THEOREM 4. *If every integer $> s$ and $\leqq s + D$ is a sum of $k - 1$ integral n-th powers $\geqq 0$, and if m is the maximum integer satisfying*

$$(m + 1)^n - m^n < D, \tag{9}$$

every integer $> s$ and $\leqq s + D + (m + 1)^n$ is a sum of k integral n-th powers $\geqq 0$.

Increasing the interval in Theorem 3 by b 18 times, we infer that 232 powers suffice from s to $s + 31b$. Since (9) holds if $m = 3$, $D = 31b$, Theorem 4 with

[4] The values of m marked * are true maxima for minimum decompositions (§10).

[5] Bull. Amer. Math. Soc., vol. 39 (1933), p. 710, Theorem 10.

$k = 233$ implies that 233 powers suffice from s to $s + 31b + c$. Add c 12 more times. Thus 245 powers suffice to s plus $31b + 13c = 234\ 578\ 479$.

Take the latter as a new D and note that $d - c < D$. Since (9) now holds if $m = 4$, Theorem 4 shows that 246 powers suffice from s to $s + D + d$ (briefly we may add d). Adding $6d$, we see that 252 powers suffice from s to

$$s + 1\ 943\ 562\ 854. \tag{10}$$

We may add f. We add $5f$, $4g$, $3h$, $2\cdot 9^{12}$, $2\cdot 10^{12}$, x^{12} ($x = 11, \cdots, 15$). We find that 273 powers suffice from s to beyond

$$L_0 = 224\ 715\ 123 \times 10^6. \tag{11}$$

We have reached the stage where single ascents are unnecessary, but may make t ascents at once by use of the writer's result (l.c., p. 711, $n = 12$):

THEOREM 5. *If all integers between s and L_0 inclusive are sums of k integral twelfth powers, then all between s and L_t inclusive are sums of $k + t$ integral powers if* $\log L_t = (12/11)^t(\log L_0 + 12 \log V) - 12 \log V$, $12V = 1 - s/L_0$.

We take $k = 273$ and L_0 as in (11), $s/L_0 = .00036957$, $\log V = -1.0793417$,

$$\log L_0 = 14.3516323, \quad \log L_0 + 12 \log V = 1.3995319,$$

$$\log \log L_t = .0377885t + .1459828. \tag{12}$$

9. **Proof of Theorems** 1, 2. In the current number of *Annals of Mathematics* the writer gives amplifications and generalizations of the remarkable paper by Vinogradow (ibid.) on the asymptotic Waring problem, and in particular obtains the following fact. If $\log \log N = 7.5068$, all integers $\geqq N$ are sums of 586 twelfth powers. By (12), $L_t > N$ if $t \geqq 195$. Hence all integers $\geqq s$ are sums of 586 twelfth powers. This can be reduced to 440.

But[6] 1560 powers suffice from $3b$ to 5×10^{27}, 2405 powers suffice from $2b$ to $3b$, and I powers from 1 to $2b$.

10. **Table** M **of minimum decompositions.** We first constructed table M, here omitted. The part with $B = 0$ required 240 equations in addition to those used for the corresponding part of our above table T. The parts with $B = 1, 2, 3$ required 45 further equations. Still further equations were used to prove that if $B = 10$, $A + m \leqq 189$, and if $B = 22$, $A + m \leqq 186$, both for every A, while 224 twelfth powers suffice from s to

$$10d + g + h = s + 8d = s + 1\ 953\ 125\ 000.$$

The discussion following (10) applies also here. Hence 245 powers suffice from s to (11) increased by 10^7. The remarkable Theorem 3 cannot be improved by use of M.

UNIVERSITY OF CHICAGO.

[6] Bull. Amer. Math. Soc., vol. 39 (1933), pp. 709, 713.

GROUPS OF CREMONA TRANSFORMATIONS IN SPACE OF PLANAR TYPE. II

By Arthur B. Coble

1. **Introduction.** We have defined in part I[1] of this account the meaning to be attached to the phrase "of planar type" and have given one example of a group G of this type. It is the purpose of this article to give further examples of groups G, which differ in some essential respects from the first.

The stable character of the elements of G is due to the fact that all of the elements have a common F-curve of the first kind. In the example given (cf. footnote 1) the elements had in addition variable isolated F-points. In the first three examples given below the elements have also an F-curve of the first kind, which may vary with the element, and whose nature is dependent upon that of the variable isolated F-points. In the fourth example given below there is also a fixed isolated F-point.

In order to ensure that the elements of G have a common F-curve of the first kind, it is convenient to define G by means of *involutorial* generators. For the first three of our groups G we use generators of types given by Sharpe and Snyder.[2] We develop anew the properties of these generators by a mapping process.

In such a provisional exploration as this, it is convenient to avoid the complications of contact singularities. Simplifying assumptions in this direction are sometimes made.

The first group G developed in §4 is considered in more detail than the later ones. Since the groups G are all associated with linear groups g generated by involutorial elements of a particular arithmetic character, it would seem preferable to discuss the groups g more generally before making applications to the Cremona groups G. This the author hopes to do in an early paper.

2. **Webs of cubic surfaces of degree two.** Since the generic surface of a web can have only fixed singularities, we distinguish two cases: (a) the generic surface has no singularities; (b) the generic surface has a node. If, in case (a), K_1 is a generic surface of the web, and $\lambda_2 K_2 + \lambda_3 K_3 + \lambda_4 K_4$ a residual net, then this net must cut K_1 in a fixed curve C and a variable net of degree two. If this net on K_1 has fixed base points, we require them to be *simple* base points. For

Received February 24, 1936.

[1] A. B. Coble, *Groups of Cremona transformations in space of planar type*, this journal, vol. 2 (1936), pp. 1–9.

[2] F. R. Sharpe and V. Snyder, *Certain types of involutorial space transformations*, Transactions of the American Mathematical Society, vol. 21 (1930), pp. 52–78.

a double base point would imply a contact for surfaces of the web, which we wish to avoid.

Let K_1 be mapped from the plane E by the system of cubic curves $(q_1 \cdots q_6)^3$ on six points q. The net of curves on K_1 is the map of a planar net $(q_1^3 \cdots q_6^3)^9$. If the net on K_1 has a fixed curve C, the net on E must decompose into a fixed curve c and a residual net of degree two. Moreover, the fixed base points of this residual net outside $q_1, \cdots, q_6$ must be simple base points if the net on K_1 has simple base points as required.

Planar nets of degree two are either the Geiser net of cubics $(1^7)^3$ or the transform of such a net by a Cremona transformation, such as $(2^2 1^6)^4$, $(2^5 1^3)^5, \cdots$. In the case of these nets of higher order, all of the multiple base points must be found among $q_1, \cdots, q_6$. Moreover, there are at least three of such multiple base points except in the case $(2^2 1^6)^4$. In this case, one of the simple base points also must be in $q_1, \cdots, q_6$. Otherwise the fixed part c of the net $(q_1^3 \cdots q_6^3)^9$ would be a $(q_1^3 \cdots q_4^3 q_5 q_6)^5$, which cannot exist. Thus the three multiple base points of highest order are always in $q_1, \cdots, q_6$, and a quadratic transformation with F-points at these three points will reduce the order of the net. The residual net can, therefore, be reduced eventually to the Geiser net $(1^7)^3$ by a process which merely shifts the mapping double six on K_1.

Thus we find just four essentially distinct degenerations of $(q_1^3 \cdots q_6^3)^9$ into a fixed curve c, and a Geiser net, namely:

	fixed curve c	*Geiser net*
I	$(q_1^2 \cdots q_6^2)^6$	$(q_1 \cdots q_6 r_1)^3$
II	$(q_1^3 q_2^2 \cdots q_6^2)^6$	$(q_2 \cdots q_6 r_1 r_2)^3$
III	$(q_1^3 q_2^3 q_3^2 \cdots q_6^2)^6$	$(q_3 \cdots q_6 r_1 r_2 r_3)^3$
IV	$(q_1^3 q_2^3 q_3^3 q_4^2 q_5^2 q_6^2)^6$	$(q_4 q_5 q_6 r_1 \cdots r_4)^3$.

The cases break off because a fixed curve c of type $(q_1^3 \cdots q_4^3 q_5^2 q_6^2)^6$ cannot exist. In these four cases the fixed curve c maps into respectively a space sextic of genus four, a space quintic of genus two, a rational space quartic, and a set of three skew lines. The points $r_1, r_2, \cdots$ map into simple base points $p_1, p_2, \cdots$ of the web.

In case (b) let the generic cubic surface K_1 have a node fixed for the entire web at N. The surface K_1 can be mapped on the plane E by cubics $(q_1 \cdots q_6)^3$, where $q_1, \cdots, q_6$ are on a conic whose points map into directions on K_1 at N. A cubic surface on N cuts K_1 in a curve whose planar map $(q_1^3 \cdots q_6^3)^9$ contains the conic $(q_1 \cdots q_6)^2$. If the surface also has a node at N, the conic appears twice in the map, and the residual curve is a $(q_1 \cdots q_6)^5$. Since the net $\lambda_2 K_2 + \lambda_3 K_3 + \lambda_4 K_4$ with node at N must meet K_1 in a net of curves with a fixed part C and a variable part of degree two; $(q_1 \cdots q_6)^5$ must have a fixed part c and a variable net of degree two, which must be the Geiser net $(1^7)^3$, since the multiplicities at q are all unity. Thus c is a conic. If c is not on

four points q, it will meet the conic $(q_1 \cdots q_6)^2$ in points not at q, whence C on K_1 will pass through N. Since we wish to avoid, as far as possible, incidences of isolated F-points with F-curves of the first kind, we take c on four points q, and obtain a fifth web for which

V: $\quad c\colon (q_1 \cdots q_4)^2;\quad$ Geiser net: $(q_5 q_6 r_1 \cdots r_5)^3$.

The conic c maps into a conic on K_1, and the points r into simple base points p of the web.

In each case the existence of the net of degree two on a surface of the web is a consequence of the existence of the corresponding net in a planar map, whence

(1) *There is a web of cubic surfaces of degree two defined by each of the following bases:*

I: *a space sextic C_6 of genus four and a simple point p_1;*
II: *a space quintic C_5 of genus two and two simple points p_1, p_2;*
III: *a rational space quartic C_4 and three simple points p_1, p_2, p_3;*
IV: *three skew lines C_3 and four simple points $p_1, \cdots, p_4$;*
V: *a conic C_2, a node N, and five simple points $p_1, \cdots, p_5$.*

Each base may be chosen in generic position in space. The net of surfaces of the web on a point x is also on x', and x, x' are correspondents under a Cremona involution I_1, I_{12}, I_{123}, I_{1234}, and I_{12345}, respectively.

There is a variety of particular cases of each of the above involutions which depends upon various degenerations of the curves C. These we do not pursue but note merely that no one of the cases I, $\cdots$, V is a particular case of another. The group G generated by involutions of type I_1 has been considered in Part I.[1] We proceed to examine the remaining involutions and the groups G generated by them.

3. The Cremona involution I_{12}. Let C_5 be a space quintic curve of genus two on a quadric Q, and let p_1, p_2 be two points in generic position with respect to C_5 but, in any case, not on Q. Then the Cremona involution I_{12} of **2** (1) is determined by the web of cubic surfaces $(C_5 p_1 p_2)^3$. The web contains surfaces $(C_5 p_1^2 p_2)^3$, $(C_5 p_1 p_2^2)^3$, which, from the definition of I_{12}, are the P-surfaces of the isolated F-points p_1, p_2, respectively.

If the generic cubic surface K_1 of the web is mapped as in **2** upon the plane E, C_5 maps into the sextic $(q_1^3 q_2^2 \cdots q_6^2)^6$, and the residual net of curves on K_1 is mapped into the Geiser net $(q_2 \cdots q_6 r_1 r_2)^3$. The Geiser octavic involution G determined by this net is the map on E of I_{12} on K_1. The members of the Geiser net with nodes at r_1, r_2, respectively, are unique. Hence the above P-surfaces are unique.

All points x on a generator of Q trisecant to C_5 are on the same net of the web K, and thus determine the same point x'. The locus of these points x' determined by a variable trisecant generator is an F-curve L of the first kind

whose P-surface is Q. Since K_1 contains one trisecant variable with K_1, then K_1 must meet L in one variable point. This trisecant is mapped on E by the directions at q_1, and these pass by G into the directions at s_1, where s_1 is the 9-th base point of the pencil $(q_2 \cdots q_6r_1r_2q_1s_1)^3$. But this is the map of the pencil on K_1 cut out by the pencil of planes on p_1, p_2, each of which with Q is a surface of the web K. Thus S_1 is the map on E of the further intersection of the line p_1p_2 with K_1, and L is the line p_1p_2. Since a plane π meets each trisecant once, the F-curve L is a simple curve on the homaloidal web H_{12} of I_{12}.

A generic plane π cuts K_1 in a cubic curve which passes by I_{12} into the section of K_1 by H, that member of H_{12} which corresponds to π under I_{12}. In E the section π is a $(q_1 \cdots q_6)^3$ which passes by G into a curve of the web $(q_2^3 \cdots q_6^3r_1^4r_2^4s_1)^9$. If we add to this planar web the fixed curve $\{(q_1^3q_2^2 \cdots q_6^2)^6\}^3$, i.e., the map of C_5 taken three times, it becomes a web $(q_1^9 \cdots q_6^9r_1^4r_2^4s_1)^{27}$, which is the map on E of the homaloidal web $H_{12} = (LC_5^3p_1^4p_2^4)^9$.

The quadric Q is a P-surface containing only the F-curve C_5. It must, therefore, pass by I_{12} into the P-surface of C_5, but on this P-surface L must have a multiplicity one greater than the normal multiplicity to indicate that Q is the P-surface of L. Hence the P-surface of C_5 is an $(L^3C_5^6p_1^8p_2^8)^{18}$.

To verify the completeness of the enumeration of isolated F-points and F-curves of the first kind, we transform $H_{12} = (LC_5^3p_1^4p_2^4)^9$ back into the web of planes by another application of I_{12}. The transform is an $(L^9C_5^{27}p_1^{36}p_2^{36})^{81}$. From this the four P-surfaces must separate, respectively, once, three times, four times, four times. There is left only $(L^0C_5^0p_1^0p_2^0)^1$, the web of planes.

We observe also that, if π_{12} is any plane on p_1, p_2, π_{12}, Q is a member of K invariant under I_{12}. Hence the variable part π_{12} of this surface is also invariant. The web K cuts π_{12} in the net of curves $(p_1p_2j_3 \cdots j_7)^3$, where $j_3, \cdots, j_7$ are the meets of π_{12} and C_5. Hence I_{12} on π_{12} is the Geiser involution G' determined by this net. Under G' the line p_1p_2 passes into the conic $(j_3 \cdots j_7)^2$, the section of Q by π_{12}. The directions on π_{12} about j_3 correspond under G' to the cubic $(p_1p_2j_3^2j_4 \cdots j_7)^3$. The five cubics of this sort, and L taken three times, make up the complete intersection of π_{12} and the P-surface of C_5. Thus the P-curve which corresponds to an F-point j on C_5 is a plane cubic in the plane $(jp_1p_2)^1$ with node at j and simple points at p_1, p_2, and the remaining four intersections of the plane with C_5.

We seek now the F-curves of the second kind of I_{12}, curves contained on every surface of $H_{12} = (LC_3^5p_1^4p_2^4)^9$. The four bisecants $l_4^{(1)}$ from p_1 to C_5 and the four $l_4^{(2)}$ from p_2 to C_5 are each F-curves of the second kind. The two P-surfaces P_{p_1}, P_{p_2} meet outside C_5 in an elliptic quartic 8-secant to C_5 with nodes at p_1 and p_2. This degenerates into two conics k_2, k_2' each on p_1, p_2 and 4-secant to C_5, and therefore in the base of H_{12}. Since these conics are cut by planes π in two points, they are double on the surfaces of H_{12}. Also the two trisecants g, g' of C_5 on the points where L cuts Q are in the base of H_{12}. For the same reason as indicated above for L, g and g' are three-fold on the P-surface of C_5.

The completeness of the above enumeration of F-curves is verified by taking from the common curve of order 81 of two surfaces H the common F-curves indicated. There is left a curve of order 9, the proper transform under I_{12} of the common line of the two planes which correspond to the surfaces H. Hence

(1) *The homaloidal web and P-surfaces of the involution I_{12} have the following description in terms of the F-system:*

$$H_{12}\,(LC_5^3p_1^4p_2^4)^9(l_4^{(1)}l_4^{(2)}k_2^2k_2'^2gg')\ ;$$

$$P_L\ = Q(C_5)^2(g,\,g')\ ;$$

$$P_{C_5} = (L^3C_5^6p_1^8p_2^8)^{18}(l_4^{(1)^2}l_4^{(2)^2}k_2^4k_2'^4g^3g'^3)\ ;$$

$$P_{p_1} = (C_5p_1^2p_2)^3(l_4^{(1)}k_2k_2')\ ;$$

$$P_{p_2} = (C_5p_1p_2^2)^3(l_4^{(2)}k_2k_2')\ .$$

It is important for the sequel to note the subordinate rôle played by the F-curve L of the first kind. Both it and its P-surface Q are determined by the choice of the other F-loci C_5, p_1, p_2.

4. **The Cremona group G generated by involutions I_{12} and its linear group g.** Let p_1, p_2, p_3, p_4 be points in generic position with respect to C_5. There are three types of products that may be formed from two involutions I_{ij} $(i,\,j = 1, \cdots, 4)$. The first type, $I_{12}I_{12}$, is the identity; the second, $I_{12}I_{13}$, has a homaloidal web of the form $(C_5^6q_1^4q_2^4q_3^8)^{15}$, where $q_1, q_3 = p_1, p_3$ and q_2 is the transform of p_2 by I_{13}; and the third, $I_{12}\,I_{34}$, has a web $(C_5^9q_1^4q_2^4q_3^{12}q_4^{12})^{27}$, where $q_3, q_4 = p_3, p_4$ and q_1, q_2 is the transform of p_1, p_2 by I_{34}.

Let G be the Cremona group generated by involutions I_{12} for variable p_1, p_2 but fixed C_5. We shall be interested mainly in the *types* of Cremona transformations in G. In forming a generic element $\Pi = \Pi'\,I_{rs}\,\Pi''$ of G, the F-points p_r, p_s of I_{rs} may fall, wholly or in part, among the isolated F-points of the web of Π', as in the very special cases above. We do not consider the case when one or two of the points p_r, p_s fall on P-surfaces of the web of Π'. Products of this latter kind have coalescent F-loci. They may be regarded as particular cases of the more general products.

The direct and inverse homaloidal webs of any product Π will each have one F-curve L_Π of the first kind whose P-surface is Q. If $\Pi = I_{rs}\cdot\Pi'$, then L_Π is the transform of the line p_rp_s by Π'; if $\Pi = \Pi''\cdot I_{rs}$, then $L_{\Pi^{-1}}$ is the transform of p_rp_s by $(\Pi'')^{-1}$. The knowledge of these F-curves is not necessary to express an element of G as a product Π.

With respect to I_{12} a linear system of surfaces has a *characteristic* $\{y\}$, i.e., an order y, a multiplicity y on L, y_0 on C_5, and y_1, y_2 on p_1, p_2. This linear system

is transformed by I_{12} into a linear system of characteristic $\{y'\}$, where $\{y'\}$, according to **3** (1), is expressed in terms of $\{y\}$ by

$$
(1)\qquad
\begin{aligned}
y' &= 9y - 2\bar{y} - 18y_0 - 3y_1 - 3y_2 ,\\
\bar{y}' &= y \qquad\quad - 3y_0 ,\\
y_0' &= 3y - \bar{y} - 6y_0 - y_1 - y_2 ,\\
y_1' &= 4y \qquad\quad - 8y_0 - 2y_1 - y_2 ,\\
y_2' &= 4y \qquad\quad - 8y_0 - y_1 - 2y_2 .
\end{aligned}
$$

This linear transformation of determinant -1 is involutorial, as we should expect. It is clear that $y' - 3y_0' = \bar{y}$, $\bar{y}' = y - 3y_0$. If we make the change of variable

$$(2)\quad y - 3y_0 = z , \qquad \bar{y} = \bar{z} , \qquad y - 2y_0 = x_0 , \qquad y_1 = 4x_1 , \qquad y_2 = 4x_2 ,$$

the linear transformation (1) interchanges z, $\bar{z}$ and yields the following transformation on x_0, x_1, x_2:

$$
(3)\qquad i_{12}:\quad
\begin{aligned}
x_0' &= 3x_0 - 4x_1 - 4x_2 ,\\
x_1' &= x_0 - 2x_1 - x_2 ,\\
x_2' &= x_0 - x_1 - 2x_2 ,\\
x_j' &= x_j \qquad\qquad (j > 2) .
\end{aligned}
$$

The last equation in (3) merely expresses that other multiplicities are unaltered.

The transformation i_{12} has period two and determinant $+1$. We observe that if our linear system has no special relation to L, i.e., if $\bar{y} = \bar{z} = 0$, then $y' - 3y_0' = z' = 0$. If also $y - 3y_0 = z = 0$, then $\bar{z}' = 0$. For the homaloidal webs mentioned above, those of I_{12}, $I_{12}I_{13}$, $I_{12}I_{34}$, $y - 3y_0$ does vanish. Moreover, when these webs are transformed by further involutions I_{ij}, the webs have no special relation to L_{ij}, the subsidiary F-curve of I_{ij}. We are therefore entitled, in transforming these webs, to set $z = \bar{z} = 0$, and then have, in the transform, $z' = \bar{z}' = 0$. Thus the significant effect of I_{ij} is that represented by i_{ij} $(i, j = 1, \cdots, \rho;\ i \neq j)$ on the variables x_0, x_i, x_j alone, and we pass from the characteristic $\{x\}$ to the characteristic $\{y\}$ as in (2). A single exception is the web of planes for which $x_0, x_1, x_2 = 1, 0, 0$ and $z, \bar{z} = 1, 0$. However, this exception disappears immediately under transformation by i_{12}. It is perhaps better to remove the exception entirely by adding to the web of planes the fixed quadric $Q = (C_5)^2$, in which case the web has the characteristic $y, \bar{y}, y_0, y_1, y_2 = 3, 0, 1, 0, 0$, which yields $z, \bar{z} = 0, 0$; $x_0, x_1, x_2 = 1, 0, 0$. This is the transform by (1) of the characteristic of the web $(LC_5^3p_1^4p_2^4)^9$ plus the F-curve L with multiplicity -1, i.e., of the web $(C_5^3p_1^4p_2^4)^9$, the homaloidal web with the L-incidence disregarded. It is obvious that the multiplicity -1 must be ascribed to L, to C_5, to p_1, and to p_2, in order that (1) may furnish the corresponding P-surface.

We now state the theorem:

(4) *Let $g(\rho)$ be the linear group generated by involutions i_{jk} $(0 < j, k \leqq \rho; j \neq k)$. If the generic element of $g(\rho)$ is*

$$\begin{aligned} x_0' &= \alpha_{00}x_0 - 4\alpha_{01}x_1 - 4\alpha_{02}x_2 - \cdots - 4\alpha_{0\rho}x_\rho\,, \\ g\colon x_l' &= \alpha_{l0}x_0 - \alpha_{l1}x_1 - \alpha_{l2}x_2 - \cdots - \alpha_{l\rho}x_\rho \qquad (l = 1, \cdots, \rho)\,, \\ x_m' &= x_m \qquad (m > \rho)\,, \end{aligned}$$

then this element represents a Cremona transformation G whose homaloidal web has the form $(L_{\mathrm{II}}C_5^{\alpha_{00}}p_1^{4\alpha_{10}} \cdots p_\rho^{4\alpha_{\rho 0}})^{3\alpha_{00}}$. The P-surface of C_5 is

$$(L_{\mathrm{II}}^3C_5^{2\alpha_{00}}p_1^{8\alpha_{10}} \cdots p_\rho^{8\alpha_{\rho 0}})^{6\alpha_{00}}\,;$$

of p_k is $(C_5^{\alpha_{0k}}p_1^{\alpha_{1k}} \cdots p_\rho^{\alpha_{\rho k}})^{3\alpha_{0k}}$ $(k = 1, \cdots, \rho)$. The F-curve L_{II} of the first kind is determined by the other F-elements, and its P-surface is Q on C_5.

For we observe that the theorem is true for $g = i_{jk}$. We find that, if it is true for g and G, it is also true for gi_{12} and GI_{12}, for $gi_{1,\rho+1}$ and $GI_{1,\rho+1}$, and for $gi_{\rho+1,\rho+2}$ and $GI_{\rho+1,\rho+2}$, the formation and comparison of these products being omitted here [cf. **5**, footnote 1]. Since g itself is merely a product of involutions i_{jk}, the proof is complete.

It is sufficient to examine the effect of (3) to see that

(5) *The group $g(\rho)$ has the linear and quadratic invariants*

$$L = x_0 - x_1 - x_2 - \cdots - x_\rho\,, \qquad Q = x_0^2 - 4x_1^2 - 4x_2^2 - \cdots - 4x_\rho^2\,.$$

For the characteristic $\alpha_{00}; \alpha_{01}, \cdots, \alpha_{0\rho}$ of the homaloidal web in (4), *L and Q take the value* 1.

A table of these characteristics which arises from the web of planes by using not more than three generators i_{jk} is as follows (the 0's not being written):

(6)	1;	7; 3111	15; 63311
	3; 11	7; 222	19; 6633
	5; 211	11; 4321	23; 87331
	9; 3311	15; 55211	27; 993311 .

On the other hand, a table, arranged according to values of $\alpha_{00} < 16$, of values satisfying $L = Q = 1$ is:

$(1; 0^\rho)$, $(3; 1^2)$, $(5; 21^2)$, $(7; 31^3)$, $(7; 2^3)$, $(9; 41^4)$, $(9; 3^21^2)$,

(7) $(11; 51^5)$, $(11; 4321)$, $(13; 61^6)$, $(13; 532^2)$, $(13; 4^231)$,

$(15; 71^7)$, $(15; 641^4)$, $(15; 63^21^2)$, $(15; 5^221^2)$, $(15; 54^21)$.

In this latter table one characteristic, $(15; 641^4)$, is not geometric. For,

transformed by i_{12}, it yields a characteristic with negative α_{10}. With respect to the possibility of reducing the order α_{00} of a characteristic we prove that

(8) *Any characteristic for which $x_0 > 1$, and which satisfies $L = Q = 1$, when so arranged that $x_1 \geqq x_2 \geqq \cdots \geqq x_\rho \geqq 0$, satisfies also the inequality* $2(x_1 + x_2) > x_0$.

For on comparing $4x_1 L = 4x_1$ with $Q = 1$, and noting that $x_1x_2 \geqq x_2^2$, etc., we have, on factoring out the non-zero $x_0 - 1$, the inequality $4x_1 \geqq x_0 + 1$. We have from $Q = 1$ that $2x_1 < x_0$. Let then

$$(\alpha) \qquad 4x_1 = x_0 + 1 + m \qquad (0 \leqq m \leqq x_0 - 2)\,.$$

If $4x_2 \geqq x_0 + 1 - m$, the theorem is proved. On the other hand, the assumption

$$(\beta) \qquad 4x_2 < x_0 + 1 - m$$

leads to a contradiction. Indeed, if we multiply $L = 1$ by $4x_2$ and compare with $Q = 1$, we find that

$$(x_0 - x_1 - 1)4x_2 \geqq x_0^2 - 1 - 4x_1^2\,.$$

On strengthening this by the use of (β), and on substituting for x_1 from (α), we get $m^2 > m(x_0 - 3)$. This, however, violates the inequality (α) when $m > 0$, and it is not satisfied when $m = 0$.

The inequality (8) permits the statement:

(9) *Any characteristic which satisfies $L = Q = 1$ can be reduced by involutions i_{jk} to one of lower x_0, and eventually to $(1; 0^\rho)$, unless in the reduction process one or more of the $x_1, \cdots, x_\rho$ become negative.*

For if we apply i_{12} to the ordered characteristic (8), the inequality $2(x_1 + x_2) > x_0$ implies $x_0' < x_0$. Hence

(10) *Any solution $\alpha_{00}; \alpha_{10}, \cdots, \alpha_{\rho 0}$ of the equations $L = Q = 1$ which satisfies the set of inequalities conjugate to $x_i \geqq 0$ $(i = 1, \cdots, \rho)$ under $g(\rho)$ (an infinite set when $\rho \geqq 4$) defines as in (4) a geometrically existent homaloidal web.*

For such a solution can be reduced to the solution $(1; 0^\rho)$ corresponding to the web of planes by successive applications of involutions i_{jk} without introducing negative numbers x_i. The corresponding Cremona involutions I_{jk}, carried out in the reverse order, yield the given homaloidal web.

With respect to the case $\rho \geqq 4$, we find that

(11) *The group $g(2)$ has the order 2; the group $g(3)$ has the order 24, and is isomorphic with the even octahedral g_{12} amplified by a reflection in a plane of symmetry on only one of the three diagonals; the group $g(\rho)$ $(\rho \geqq 4)$ is of infinite order.*

Let A_k denote the set of four forms, one of which is

$$(2k^2 + k + 1)x_0 - 2(k^2 + k + 1)x_1 - 2k(k + 1)x_2 - (2k^2 + 1)x_3 - 2k^2x_4\,,$$

the other three arising from this by interchanging x_1, x_2, or also x_3, x_4. Let B_k denote the similar set of four forms, one of which is

$$(2k^2 - k + 1)x_0 - 2(k^2 - k + 1)x_1 - 2k(k - 1)x_2 - (2k^2 + 1)x_3 - 2k^2x_4 .$$

We find that i_{34} carries A_k into B_{k+1}, and that i_{12} carries B_k into A_k. Thus $i_{34}i_{12}$ sends A_k into A_{k+1}. Hence $i_{34}i_{12}$ is of infinite order and $g(4)$ is infinite. Necessarily then $g(\rho)$ $(\rho > 4)$ is also infinite.

We prove finally that

(12) *The inverse of the element g in* (4) *is obtained from g by interchanging α_{ij} and α_{ji}. The coefficients of g and g^{-1} satisfy the following relations in which the indices run from* 1 *to* ρ:

$$\begin{array}{llll}
\sum_i \alpha_{i0} = \alpha_{00} - 1 , & & \sum_j \alpha_{0j} = \alpha_{00} - 1 , & \\
\sum_i \alpha_{ij} = 4\alpha_{0j} - 1 , & & \sum_j \alpha_{ij} = 4\alpha_{i0} - 1 , & \\
4\sum_i \alpha_{i\,0}^2 = \alpha_{0\,0}^2 - 1 , & & 4\sum_j \alpha_{0\,j}^2 = \alpha_{0\,0}^2 - 1 , & \\
\sum_i \alpha_{i\,j}^2 = 4\alpha_{0\,j}^2 + 1 , & & \sum_j \alpha_{i\,j}^2 = 4\alpha_{i\,0}^2 + 1 , & \\
\sum_i \alpha_{i0}\alpha_{ij} = \alpha_{00}\alpha_{0j} , & & \sum_j \alpha_{0j}\alpha_{ij} = \alpha_{00}\alpha_{i0} , & \\
\sum_i \alpha_{ij}\alpha_{ik} = 4\alpha_{0j}\alpha_{0k} & (j \neq k) , & \sum_j \alpha_{ij}\alpha_{kj} = 4\alpha_{i0}\alpha_{k0} & (i \neq k) .
\end{array}$$

For the first column of relations expresses the invariance of L and Q under g. If in g we interchange α_{ij} and α_{ji} to obtain g', this first column of relations also expresses that $gg' = 1$; hence $g' = g^{-1}$. The second column then expresses that Q, L are invariant under g^{-1}.

5. **The Cremona involution I_{123} and its attached Cremona group G and linear group g.** Let C_4 be a rational space quartic curve and p_1, p_2, p_3 three points in generic position. The web of cubic surfaces K, which defines I_{123} as in **2** (1), contains only one degenerate member, $\pi_{123} - Q$, π_{123} being the plane $(p_1p_2p_3)^1$, and Q being the quadric on C_4. The mapping of K_1 on E by the system $(q_1 \cdots q_6)^3$ described in **2** carries C_4 into $(q_1^3q_2^3q_3^2 \cdots q_6^2)^6$, $p_1p_2p_3$ into $r_1r_2r_3$, and the net residual to C_4 cut out on K_1 by K into the Geiser net $(q_3 \cdots q_6\, r_1r_2r_3)^3$. This net determines the Geiser involution G, the map on E of I_{123} on K_1.

As in **3** we find that p_1, p_2, p_3 are isolated F-points of I_{123}. The P-surface of p_i, P_{p_i} is the surface of the web K with node at p_i $(C_4\, p_i^2\, p_j\, p_k)$ $(i, j, k = 1, 2, 3)$.

A generator of Q trisecant to C_4 corresponds to a single point x', and the locus of these points x' is an F-curve of the first kind, L, whose P-surface is Q. Since K_1 meets Q in two such variable trisecants, K_1 meets L in two variable points. The surface $K = \pi_{123}\cdot Q$ is invariant under I_{123}, and Q being a P-surface, π_{123} is itself invariant. Surfaces K meet π_{123} in the net of cubics

$(p_1p_2p_3t_4 \cdots t_7)^3$, where $t_4, \cdots, t_7$ are the meets of π_{123} and C_4. The Geiser involution G' of this net is I_{123} on π_{123}. The trisecants of C_4 cut π_{123} in the conic (π_{123}, Q). This conic is carried by G' into L, a rational quartic on π_{123} with nodes at p_1, p_2, p_3 and simple points at $t_4, \cdots, t_7$. This curve L is contained simply on the homaloidal web H_{123} of I_{123}.

The plane sections of K_1 map into $(q_1 \cdots q_6)^3$ which passes by G into the system $(q_3^4 \cdots q_6^4 r_1^5 r_2^5 r_3^5 s_1 s_2)^{12}$, s_1, s_2 being the transforms of q_1, q_2 by G. If we add to this the map of the fixed curve C_4 four times, we get a system $(q_1^{12} \cdots q_6^{12} r_1^5 \cdots r_3^5 s_1 s_2)^{36}$, the map of the curves cut out on K_1 by the homaloidal web, $H_{123} = (LC_4^4 p_1^5 p_2^5 p_3^5)^{12}$. The points s_1, s_2 arise from the two variable intersections of L and K_1. As in **3** we conclude that the P-surface of C_4 is an $(L^3 C_4^8 p_1^{10} p_2^{10} p_3^{10})^{24}$, and can again verify that H_{123} is transformed by I_{123} back into the web of planes.

We list the following F-curves of the second kind, each contained on every member of H_{123}: (a) $g_1, \cdots, g_4$, the four trisecants of C_4 from points where L meets Q outside C_4; (b) $l_3^{(1)}, l_3^{(2)}, l_3^{(3)}$, the three bisecants of C_4 from p_1, p_2, p_3 respectively; (c) c_3, the twisted cubic on p_1, p_2, p_3 and 6-secant to C_4; and (d) $k_2^{(12)}, k_2^{(13)}, k_2^{(23)}$, respectively conics on p_1, p_2; p_1, p_3; p_2, p_3 each 4-secant to C_4. Since c_3 is on $P_{p_1}, P_{p_2}, P_{p_3}$, there can be only one such curve, a triple curve on H_{123}. Since P_{p_1} and P_{p_2} meet outside C_4 in c_3 and $k_2^{(12)}$, there is but one such conic, a double curve on H_{123}. The completeness of this list can be verified as in **3**. Hence

(1) *The homaloidal web and P-surfaces of the involution I_{123} have the following description in terms of the F-system*:

$$H_{123}(LC_4^4 p_1^5 p_2^5 p_3^5)^{12}(g_1 \cdots g_4 l_3^{(1)} \cdots l_3^{(3)} k_2^{(12)^2} \cdots k_2^{(23)^2} c_3^3);$$

$$P_L = Q(C_4)^2(g_1 \cdots g_4);$$

$$P_{C_4} = (L^3 C_4^8 p_1^{10} p_2^{10} p_3^{10})^{24}(g_1^3 \cdots g_4^3 l_3^{(1)^2} \cdots l_3^{(3)^2} k_2^{(12)^4} \cdots k_2^{(23)^4} c_3^6);$$

$$P_{p_i} = (C_4 p_i^2 p_j p_k)^3 (l_3^{(i)} c_3 k_2^{(ij)} k_2^{(ik)}) \qquad (i, j, k = 1, 2, 3).$$

To an F-point on C_4 there corresponds a P-curve on P_{C_4}, which is a rational space quartic on p_1, p_2, p_3, cutting L once, and 8-secant to C_4. This is a consequence of the multiplicites of C_4 on H, on P_L, on P_{p_i}, and on P_{C_4}. We observe that a curve of this order and these multiplicities must be contained on P_{C_4}.

We consider the Cremona group G generated by involutions I_{123} for variable p_1, p_2, p_3 but fixed C_4. The same remarks as are made in **4** with respect to products of these involutions and with respect to the behavior of L apply here. Let $\{y\}$ again be the characteristic of a linear system of surfaces with respect to the F-basis of H_{123}. The linear transformation of this characteristic produced by I_{123} is easily written by using (1). If we make the following change of variable in this characteristic:

(2) $\quad y - 3y_0 = z, \quad \bar{y} = \bar{z}, \quad y - 2y_0 = x_0, \quad y_i = 5x_i \quad (i = 1, 2, 3),$

then, apart from $z' = \bar{z}$, $\bar{z}' = z$, the transformation of the characteristic is expressed by

$$
(3) \qquad i_{123}: \begin{aligned} x_0' &= 4x_0 - 5x_1 - 5x_2 - 5x_3, \\ x_i' &= x_0 - x_1 - x_2 - x_3 - x_i \qquad (i = 1, 2, 3), \\ x_j' &= x_j \qquad (j > 3). \end{aligned}
$$

This involution, i_{123}, of determinant -1 has the invariant linear and quadratic forms:

$$
(4) \quad L = x_0 - x_1 - x_2 - \cdots - x_\rho, \qquad Q = x_0^2 - 5x_1^2 - 5x_2^2 - \cdots - 5x_\rho^2.
$$

If the linear group generated by involutions i_{jkl} $(j, k, l = 1, \cdots, \rho)$ has the generic element

$$
(5) \qquad g: \begin{aligned} x_0' &= \alpha_{00}x_0 - 5\alpha_{01}x_1 - 5\alpha_{02}x_2 - \cdots - 5\alpha_{0\rho}x_\rho, \\ x_l' &= \alpha_{l0}x_0 - \alpha_{l1}x_1 - \alpha_{l2}x_2 - \cdots - \alpha_{l\rho}x_\rho \qquad (l = 1, \cdots, \rho), \\ x_m' &= x_m \qquad (m > \rho), \end{aligned}
$$

we prove as before that this element represents a Cremona transformation in G with a homaloidal web $(L_\tau C_4^{\alpha_{00}} p_1^{5\alpha_{10}} \cdots p_\rho^{5\alpha_{\rho 0}})^{3\alpha_{00}}$.

This linear group g becomes infinite when $\rho \geqq 5$. An easy verification of this is obtained from two sets of values of $x_0; x_1, x_2, \cdots, x_5$, namely:

$$
A_k\colon 5(k^2 - k + 1);\, k^2, k^2, k^2 - k + 1, (k-1)^2, (k-1)^2;
$$

$$
B_k\colon 5(k^2 - k + 1);\, (k-1)^2, (k-1)^2, k^2 - k + 1, k^2, k^2.
$$

Since i_{345} sends A_k into B_{k+1}, and i_{123} sends B_k into A_{k+1}, then $i_{345}i_{123}$ sends A_k into A_{k+2}, and thus has an infinite period.

6. **The Cremona involution I_{1234}; its Cremona group G and linear group g.** Let C_3 be three skew lines on a quadric Q, and $p_1, \cdots, p_4$ four points in generic position. Then I_{1234} is defined as in **2**(1) by the web $K = (C_3 p_1 \cdots p_4)^3$. If K_1 is mapped upon E by $(q_1 \cdots q_6)^3$, C_3 mapping into $(q_1^3 q_2^3 q_3^3 q_4^2 q_5^2 q_6^2)^6$ and $p_1 \cdots p_4$ into $r_1 \cdots r_4$, the net residual to C_3 cut out on K_1 by K maps into the Geiser net $(q_4 q_5 q_6 r_1 \cdots r_4)^3$, which determines the involution G, the map on E of I_{1234} on K_1. As before, the surfaces of K with nodes at $p_1, \cdots, p_4$, respectively, are the P-surfaces of the isolated F-points $p_1, \cdots, p_4$. Also Q is the P-surface of an F-curve of the first kind L contained simply on the homaloidal web H_{1234} of I_{1234}, and met in three variable points by a surface K.

The web $(q_1 \cdots q_6)^3$ is transformed by G into a web $(q_4^5 q_5^5 q_6^5 r_1^6 \cdots r_4^6 s_1 s_2 s_3)^{15}$, where s_i is the transform of q_i $(i = 1, 2, 3)$. If to this we add the map of C_3 five times, we get a system $(q_1^{15} \cdots q_6^{15} r_1^6 \cdots r_4^6 s_1 \cdots s_3)^{45}$, whence $H_{1234} = (LC_3^5 p_1^6 \cdots p_4^6)^{15}$.

A member of H_{1234} cuts Q in 15 cross-generators, whence the order of L is 15. The P-surface of p_1 cuts Q in 3 cross-generators, whence L has triple points at

$p_1, \cdots, p_4$, the three directions at p_1 on L not being on P_{p_1}, since these latter directions are self-corresponding.

Let the three skew lines C_3 be $\lambda_1, \lambda_2, \lambda_3$. The web K contains the degenerate member $(\lambda_1 p_1)^1 \cdot (\lambda_2\lambda_3 p_2 p_3 p_4)^2$, invariant under I_{1234}. The plane $(\lambda_1 p_1)^1$ is cut by the web K in a net of curves of degree one only, whence I_{1234} transforms the plane into the quadric. On taking away this quadric, and the surface P_{p_1}, from a member of H_{1234}, we have a surface $(L\lambda_1^4\lambda_2^3\lambda_3^3 p_1^4 \cdots p_4^4)^{10}$, the P-surface of λ_1. This cuts Q in 10 cross-generators, whence L is 10-secant to λ_1. Thus L, of order 15 with triple points at $p_1, \cdots, p_4$ and 10-secant to each of $\lambda_1, \lambda_2, \lambda_3$, is met by a member of K in 3 variable points as expected.

The F-curves of the second kind of H_{1234} consist of 12 lines $l_{i,j}$ from point p_i across the two lines $\lambda_k\lambda_l$ $(i = 1, \cdots, 4; j, k, l = 1, 2, 3)$; and of four cubic curves $c_3^{(i)}$ on p_j, p_k, p_l and bisecant to each of the three lines λ $(i, \cdots, l = 1, \cdots, 4)$, these curves being three-fold on H_{1234}. On eliminating the F-curves of both kinds from the common curve of two members of H_{1234}, we have a curve of order 15, the proper transform of a line under I_{1234}. Hence

(1) *The homaloidal web and P-surfaces of the involution I_{1234} have the following description in terms of the F-system*:

$$H_{1234}(LC_3^5 p_1^6 \cdots p_4^6)^{15}(c_3^{(1)^3} \cdots c_3^{(4)^3} l_{1,1} \cdots l_{4,3});$$

$$P_L = Q(C_3)^2;$$

$$P_{\lambda_1} = (L\lambda_1^4\lambda_2^3\lambda_3^3 p_1^4 \cdots p_4^4)^{10}(c_3^{(1)^2} \cdots c_3^{(4)^2} l_{1,2} \cdots l_{4,2} l_{1,3} \cdots l_{4,3});$$

$$P_{p_1} = (C_3 p_1^2 p_2 p_3 p_4)^3 (c_3^{(2)} c_3^{(3)} c_3^{(4)} l_{1,1} l_{1,2} l_{1,3}).$$

To an F-point on λ_1 there corresponds a P-curve on P_{λ_1}, which is a rational space quintic on $p_1, \cdots, p_4$, 1-secant to L, 4-secant to λ_1, and 3-secant to each of λ_2, λ_3.

The involutions I_{1234} with fixed triad of lines C_3 and variable $p_1, \cdots, p_4$ generate a Cremona group G of ternary type. With the same change of characteristic as before (except that we set $y_i = 6x_i$ $(i = 1, \cdots, 4)$) we find the following expression of the effect of I_{1234}:

$$\begin{aligned} & x_0' = 5x_0 - 6x_1 - 6x_2 - 6x_3 - 6x_4, \\ (2) \qquad i_{1234}\colon\ & x_i' = x_0 - x_1 - x_2 - x_3 - x_4 - x_i \qquad (i = 1, \cdots, 4), \\ & x_j' = x_j \qquad (j > 4). \end{aligned}$$

This involution i_{1234} of determinant $+1$ has the invariant linear and quadratic forms

$$(3) \quad L = x_0 - x_1 - x_2 - \cdots - x_\rho, \qquad Q = x_0^2 - 6x_1^2 - 6x_2^2 - \cdots - 6x_\rho^2.$$

In this case the form of the generic element of the group g generated by elements i_{klmn} $(k, \cdots, n = 1, \cdots, \rho)$ is

$$
(4) \qquad g: \begin{aligned} x_0' &= \alpha_{00}x_0 - 6\alpha_{01}x_1 - \cdots - 6\alpha_{0\rho}x_\rho, \\ x_l' &= \alpha_{l0}x_0 - \alpha_{l1}x_1 - \cdots - \alpha_{l\rho}x_\rho \qquad (l = 1, \cdots, \rho), \\ x_m' &= x_m \qquad (m > \rho). \end{aligned}
$$

To this element g of the linear group there corresponds an element G of the Cremona group with a homaloidal web of the form $(L_r C_3^{\alpha_{00}} p_1^{6\alpha_{01}} \cdots p_\rho^{6\alpha_{0\rho}})^{3\alpha_{00}}$.

The linear group g is infinite if $\rho \geqq 6$. For if

$$A_k = (k^2 + k + 1)(x_0 - x_3 - x_4) - (k + 1)^2(x_1 + x_2) - k^2(x_5 + x_6),$$

$$B_k = (k^2 + k + 1)(x_0 - x_3 - x_4) - k^2(x_1 + x_2) - (k + 1)^2(x_5 + x_6),$$

then i_{3456} transforms A_k into B_{k+1}, and i_{1234} transforms B_{k+1} into A_{k+2}, whence $i_{3456}i_{1234}$ has an infinite period.

7. **The Cremona involution I_{12345}; its Cremona group G and linear group g.** Let I_{12345} be the involution defined as in **2**(1) by the web $K = (C_2N^2p_1 \cdots p_5)^3$, and let $Q = (N^2C_2)^2$ be the quadric cone with node at N and on the conic C_2. The generators of Q are P-curves whose F-points run over an F-curve of the first kind L contained simply on the homaloidal web $H_{1\cdots5}$ of $I_{1\cdots5}$.

The sections of K_1 by the web K yield, under the mapping of **2**(b), the Geiser net $(q_5q_6r_1 \cdots r_5)^3$ with involution G. Plane sections of K_1 map into the web $(q_1 \cdots q_6)^3$ on E, and this passes by G into the web $(q_5^6q_6^6r_1^7 \cdots r_5^7s_1 \cdots s_4)^{18}$, the partial map of the sections of K_1 by $H_{1\cdots5}$, $s_1, \cdots, s_4$ being the transforms of $q_1, \cdots, q_4$ by G. We add to this partial map the conic $\{(q_1 \cdots q_4)^2\}^6$, taken six times, in order to have like multiplicities for each q. We add also the conic $\{(q_1 \cdots q_6)^2\}^{12}$, taken 12 times, in order to bring the multiplicities of the q's up to one third of the order of the web. The resulting web of order 54 shows that $H_{1\cdots5} = (LC_2^6N^{12}p_1^7 \cdots p_5^7)^{18}$. The simple points $s_1, \cdots, s_4$ above arise from the four variable intersections of K_1 and L. As before, the P-surface of p_1 is $P_{p_1} = (C_2N^2p_1^2p_2 \cdots p_5)^3$. These P-surfaces, and all surfaces K, meet Q outside C_2 in four generators. A member of $H_{1\cdots5}$ meets Q in 24 variable generators, in addition to C_2, whence L has the order 24.

The web K contains the degenerate member $(C_2)^1\cdot(N^2p_1 \cdots p_5)^2$. Since the plane is not cut by K in a net of degree two, the plane and quadric interchange under $I_{1\cdots5}$, the P-surface of C_2 separating from the homaloid which corresponds to the plane. Hence the P-surface of C_2 is an $(LC_2^6N^{10}p_1^6 \cdots p_5^6)^{10}$. Moreover, the quadric passes into the plane. On taking this plane and the surfaces $P_{p_1}, \cdots, P_{p_5}$ from twice a homaloid we have twice the P-surface of N. Thus the P-surface of N is an $(LC_2^3N^7p_1^4 \cdots p_5^4)^{10}$.

The P-surfaces of C_2 and N meet Q outside C_2 in 20 and 14 generators respectively, whence L is 20-secant to C_2 and has a 14-fold point at N. These multiplicities account for the four variable intersections of L and the web K. We can verify as usual that $H_{1\cdots5}$ is transformed by $I_{1\cdots5}$ back into the web of planes.

The F-curves of the second kind include the five lines $l^{(i)}$ from N to p_i

$(i = 1, \cdots, 5)$ and the 10 conics $c_2^{(ij)}$ on N, p_i, p_j and the two points on C_2 cut out by the plane $(Np_ip_j)^1$ $(i, j = 1, \cdots, 5;\ i \neq j)$. Another F-curve appears from the following considerations. Surfaces K cut the plane $(C_2)^1$ in the net of straight lines, and the cone Q in a net of sextic curves with a 4-fold point at N, on $p_1, \cdots, p_5$, and 4-secant to C_2. For example K_1, mapped on E, yields the line q_5q_6 and the conic $(r_1 \cdots r_5)^2$. This conic maps into the sextic curve mentioned. On the particular surface $K = P_{p_1}$, this sextic consists of the generator Np_1 and a residual quintic curve c_5 with a triple point at N on $p_1, \cdots, p_5$ and 4-secant to C_2. Obviously c_5 is on every P_{p_i}, and also on $H_{1\cdots5}$, necessarily 5-fold. We verify now that two surfaces of $H_{1\cdots5}$ meet outside the F-curves in a curve of order 18, the transform of a line by $I_{1\cdots5}$. Hence

(1) *The homaloidal web and P-surfaces of the involution $I_{1\cdots5}$ have the following description in terms of the F-system:*

$$H_{1\cdots5}(LC_2^6N^{12}p_1^7 \cdots p_5^7)^{18}(l^{(1)} \cdots l^{(5)}c_2^{(12)^2} \cdots c_2^{(45)^2}c_5^5);$$

$$P_L = Q(N^2C_2)^2;$$

$$P_{C_2} = (LC_2^6N^{10}p_1^6 \cdots p_5^6)^{16}(c_2^{(12)^2} \cdots c_2^{(45)^2}c_5^4);$$

$$P_N = (LC_2^3N^7p_1^4 \cdots p_5^4)^{10}(l^{(1)} \cdots l^{(5)}c_2^{(12)} \cdots c_2^{(45)}c_5^3);$$

$$P_{p_1} = (C_2N^2p_1^2p_2 \cdots p_5)^3(l^{(1)}c_2^{(12)} \cdots c_2^{(15)}c_5).$$

To any F-point on C_2 there corresponds a P-curve on P_{C_2}, which is a sextic curve with a triple point at N, simple points at $p_1, \cdots, p_5$, cutting C_2 and L six times and once, respectively.

The elements of the three types of Cremona groups G just discussed, with fixed F-curves C_5, C_4, C_3, respectively, have had an F-curve of the first kind, L_{Π}, which was variable with the element, and which was determined by the other F-elements. Thus the element Π of G could be suitably described without a particular description of this curve L_{Π}. An essential feature of this situation is that L_{Π} has a P-surface $Q = (C_i)^2$, which is the same for all the elements of G.

In the case of the involution $I_{1\cdots5}$ above, the P-surface of L is the quadric cone with node at the isolated F-point N. In order to keep this P-surface fixed, we consider the Cremona group G generated by involutions I_{12345} *with fixed C_2 and N*, and with variable positions of $p_1, \cdots, p_5$ only.

We examine first the effect of I_{12345} upon linear systems of surfaces. Let y be the order, $\bar{y}$, y_0, s, $y_1, \cdots, y_5$ the multiplicities of L, C_2, N, $p_1, \cdots, p_5$ for such a system. After transformation by I_{12345} we obtain the new characteristic

$$
\begin{aligned}
y' &= 18y - 2\bar{y} - 16y_0 - 10s - 3y_1 - \cdots - 3y_5,\\
\bar{y}' &= y - y_0 - s,\\
y_0' &= 6y - \bar{y} - 6y_0 - 3s - y_1 - \cdots - y_5, \qquad (2)\\
s' &= 12y - 2\bar{y} - 10y_0 - 7s - 2y_1 - \cdots - 2y_5,\\
y_i' &= 7y - 6y_0 - 4s - y_1 - \cdots - y_5 - y_i \qquad (i = 1, \cdots, 5).
\end{aligned}
$$

If we set $y - y_0 - s = z$, $\bar{y} = \bar{z}$, we observe that

$$z' = \bar{z}, \qquad \bar{z}' = z, \qquad s' - 2y_0' = -(s - 2y_0).$$

The remaining equations (2) are then covered by

$$y' - 2y_0' = 6(y - 2y_0) - 4(s - 2y_0) - y_1 - \cdots - y_5,$$
$$y_i' = 7(y - 2y_0) - 4(s - 2y_0) - y_1 - \cdots - y_5 - y_i \qquad (i = 1, \cdots, 5).$$

We shall be interested only in the transforms of the web of planes for which $s - 2y_0 = 0$, and therefore $s' - 2y_0' = 0$. On dropping this term and setting

$$y - 2y_0 = x_0, \qquad y_i = 7x_i \qquad (i = 1, \cdots, 5),$$

we obtain

$$\begin{aligned} &x_0' = 6x_0 - 7x_1 - 7x_2 - \cdots - 7x_5, \\ (3) \qquad i_{12345}\colon\ &x_i' = x_0 - x_1 - x_2 - \cdots - x_5 - x_i \qquad (i = 1, \cdots, 5), \\ &x_j' = x_j \qquad (j > 5). \end{aligned}$$

This involution i_{12345} has the determinant -1 and the invariant linear and quadratic forms

$$(4) \quad L = x_0 - x_1 - x_2 - \cdots - x_\rho, \qquad Q = x_0^2 - 7x_1^2 - 7x_2^2 - \cdots - 7x_\rho^2.$$

If the generic element of the linear group g generated by i_{jklmn} $(k, \cdots, n \leqq \rho)$ has the form

$$\begin{aligned} &x_0' = \alpha_{00}x_0 - 7\alpha_{01}x_1 - \cdots - 7\alpha_{0\rho}x_\rho, \\ (5) \qquad g\colon\ &x_l' = \alpha_{l0}x_0 - \alpha_{l1}x_1 - \cdots - \alpha_{l\rho}x_\rho \qquad (l = 1, \cdots, \rho), \\ &x_m' = x_m \qquad (m > \rho), \end{aligned}$$

there corresponds to it an element G of the Cremona group with a homaloidal web of the form

$$(L_{\text{II}}\, C_2^{\alpha_{00}} N^{2\alpha_{00}} p_1^{7\alpha_{10}} \cdots p_\rho^{7\alpha_{\rho0}})^{3\alpha_{00}}.$$

We find that this linear group g is infinite when $\rho \geqq 7$. For if

$$A_k = (k^2 + k + 1)(x_0 - x_3 - x_4 - x_5) - (k + 1)^2(x_1 + x_2) - k^2(x_5 + x_6),$$
$$B_k = (k^2 + k + 1)(x_0 - x_3 - x_4 - x_5) - k^2(x_1 + x_2) - (k + 1)^2(x_5 + x_6),$$

then i_{12345} transforms B_k into A_{k+1}, and i_{34567} transforms A_k into B_{k+1}, whence $i_{12345}\, i_{34567}$ transforms B_k into B_{k+2}, and thus has an infinite period.

UNIVERSITY OF ILLINOIS.

CRITICAL POINT THEORY UNDER GENERAL BOUNDARY CONDITIONS

By Marston Morse and George B. Van Schaack

1. **Introduction.** M. Morse has previously treated the theory of the critical points of a function of class C^2, whose critical values are isolated and the neighborhoods of whose critical sets admit a special type of deformation, ref. [7, 8]. In the present paper it is assumed merely that the function f has continuous first partial derivatives which satisfy Lipschitz conditions and that the critical values (not critical points) of f are isolated. In spite of the fact that the critical sets may *not be locally connected or possess neighborhoods which are contractible,* it is shown that the type numbers of the critical sets are *finite* and depend on the definition of f only in an arbitrarily small neighborhood of the critical set. We emphasize the fact that this part of the treatment *does not depend at all upon the definition of f in the large.*

The authors have also taken up the theory of critical points on an abstract metric space, ref. [10]. The present treatment, although less general, is simpler and more suitable for most applications in analysis. The reader may also refer to the papers of A. B. Brown [2], Lefschetz [5], and Birkhoff and Hestenes [1]. See [12] for a later abstract by Morse.

The second part of this paper contains the first treatment under general boundary conditions. The principal theorem here was announced by Morse in [6]. This part of the paper has important applications in the theory of harmonic functions, as will appear in a subsequent paper by Morse.

Finally, various group theory aspects of the problem are brought out.

I. Type numbers

2. **The function and the critical set.** Let $(x) = (x_1, \cdots, x_n)$ be the rectangular coördinates of a point in euclidean n-space. Let R be a limited, open, n-dimensional region of this space. Let $f(x)$ be a real, single-valued function of[1] class C^1L defined on R. A point of R at which all of the first partial derivatives of f vanish will be called a *critical point* of f. The value of f at a critical point will be called a *critical value* of f. We assume that the critical values of f are isolated. Neighboring each ordinary (non-critical) point of f the differential equations of the trajectories orthogonal to the loci $f =$ constant may be given the form

Received February 26, 1936.

[1] A function defined on an open region will be said to be of class C^1L on this region if it is of class C^1 and if its partial derivatives satisfy Lipschitz conditions neighboring each point of the region.

$$\frac{dx_i}{dt} = f_{x_i} \qquad (i = 1, \cdots, n).$$

Since f is of class C^1L on R, there is a unique trajectory through each ordinary point on R.

By a *critical set* σ of f will be understood any closed[2] set of critical points on which f is a constant c and which is at a positive distance from all other critical points of f. A critical set may or may not be connected. In general it will not be a finite complex. By a *neighborhood* of σ we mean an open set of points of R which includes all points of R within a small positive distance of σ. We shall *admit* only those neighborhoods of σ whose closures contain no critical points at which $f \geqq c$ other than the critical points of σ. A neighborhood of σ will be termed arbitrarily small if all of its points lie within an arbitrarily small distance of σ.

3. **Deformations and maximal sets.** In §§3–7 we shall be concerned with a single critical set σ of f on R on which $f = c$. We begin by considering two deformations of R. We shall make use of the trajectories

$$\frac{dx_i}{dt} = -f_{x_i} \qquad (i = 1, \cdots, n),$$

orthogonal to the manifolds $f =$ constant. We make the convention that there is a trajectory coincident with each critical point at all times t.

The deformation $D(t)$. Let p be a point of R and let p_t be the point corresponding to $t \geqq 0$ on the trajectory which issues from p when $t = 0$. It is understood that p_t does not necessarily exist for all values of t. Under the deformation $D(t)$ the point p shall be replaced by the point p_t at the time t. The value of f at p_t is a non-increasing function of t at all values of t for which p_t is defined. Critical points are held fast under $D(t)$.

The deformation $\Delta(t)$. Recall that c is a critical value of f. Let p be a point of R at which $f > c$, and λ the trajectory which issues from p when $t = 0$. Under the deformation $\Delta(t)$, p shall be deformed as under $D(t)$, provided f remains greater than c on λ. If $f = c$ at a point p_τ on λ, p shall be deformed as under $D(t)$ for $t \leqq \tau$, and shall be replaced by p_τ for all values of $t > \tau$. The points of R at which $f \leqq c$ shall remain fixed under $\Delta(t)$. We distinguish between a deformation such as $\Delta(t)$ and the *final* image P_t of a point P under $\Delta(t)$. The deformation $\Delta(t)$ shall refer to the set of all images P_τ of P for which $0 \leqq \tau \leqq t$, while the final image of P under $\Delta(t)$ shall refer to the image P_t.

We introduce the following definition.

A neighborhood Y of σ will be said to be *Δ-contractible* if there exists a neighborhood X of σ such that the images of Y under $\Delta(t)$ lie on a closed subdomain of X for all values of $t \geqq 0$. We term Y Δ-contractible on X. This type of contractibility does not imply that Y can be deformed on X into an arbitrarily small neighborhood of σ.

[2] Closures shall be taken relative to the entire space (x).

Let VW be an arbitrary pair of neighborhoods of σ of which $W \subset V$. We shall distinguish between two types of cycles[3] neighboring σ. We shall refer to these cycles as *belonging* to σ. We shall say that a point of R at which $f < c$ is *below* c.

By a *spannable k-cycle corresponding to VW* we shall mean a k-cycle on W, below c, ~ 0 on W, but $\not\approx 0$ on V below c.

By a *critical k-cycle corresponding to VW* we shall mean a k-cycle on W, $\not\approx$ on V to a k-cycle on V below c.

Null cycles are naturally admitted, and we make the convention that on a null cycle the value of f may be chosen at pleasure, and in particular may be taken below c. By virtue of this convention no cycle which is dependent on V can be a critical cycle corresponding to VW.

We introduce the following definition.

Let the cycles of a class C of k-cycles be distinguished by the possession of certain properties. By a *maximal* set A of cycles of C will be meant a finite set of cycles of C, every proper linear combination (always mod π) of whose cycles belongs to C, and which is such that there exists no set of cycles of C, every proper linear combination of whose cycles belongs to C and which contains A as a proper subset.

The following theorem is of fundamental importance.

NEIGHBORHOOD THEOREM. *There exists a fixed neighborhood N^* of σ, and corresponding to any neighborhood X of σ on N^* a neighborhood $M(X)$ of σ, Δ-contractible on X and with the following property. Corresponding to any two pairs of neighborhoods X $M(X)$ and Y $M(Y)$ of σ, of which X and Y lie on N^*, there exist on any arbitrarily small neighborhood of σ common maximal sets of spannable and critical k-cycles.*

In §§4–6 we shall be occupied with the proof of this theorem.

The maximal sets of the theorem are of importance in that they depend only on the neighborhood of σ. We shall subsequently define type numbers of σ with their aid.

4. α-admissible neighborhoods.

We begin the proof of the Neighborhood Theorem with the following lemma.

LEMMA 4.1. *Corresponding to an arbitrary neighborhood X of σ there exists a neighborhood $A(X)$ of σ which is Δ-contractible on X.*

Let Y be a neighborhood of σ whose closure lies on X. If the lemma is false there must exist an infinite set of points P_n $(n = 1, 2, \cdots)$ which tend to a point on σ as n becomes infinite and which under $\Delta(t)$ possess images Q_n respectively on the boundary of Y. Let Q be a cluster point of the points Q_n.

[3] We deal with the case mod π, where π is any prime > 1. When the restriction mod π is omitted, the development holds for the absolute case as well. The notation $w \approx 0$ is used to imply that $nw \sim 0$, where n is an integer $\not\equiv 0$ mod π, Lefschetz [4]. More generally, the whole theory is valid for coefficients in an arbitrary commutative field, chains, cycles, and homologies being defined with coefficients in that field.

It is clear that $f = c$ at Q. The trajectory which passes through Q has a positive distance from σ, as do the trajectories which pass through points Q_n sufficiently near Q. This is contrary to the nature of the points P_n, which lie on trajectories through the respective points Q_n. Hence the lemma holds as stated.

The deformation $\Delta(t, \tau)$. We introduce the product deformation

$$\Delta(t, \tau) = \Delta(t) \cdot D(\tau) \qquad (t \geqq 0, \tau \geqq 0).$$

We understand thereby that an arbitrary point P on R is deformed into a final image $P_{t,\tau}$ as follows. The point P is first deformed under $\Delta(t)$ into a point P_t. We take $P_{t,\tau}$ as the final image of P_t under $D(\tau)$. This deformation deforms points below c through points below c.

If X is an arbitrary set of points on R, the subset of points of X below c will be denoted by X_c.

We come to the following lemma.

Lemma 4.2. *Corresponding to arbitrary neighborhoods X and N of σ, any sufficiently large value of $t \geqq 0$ and sufficiently small value of $\tau > 0$ will have the following property. Under $\Delta(t, \tau)$, $A(X)$ is deformed on X onto $N + X_c$.*

Let M be a neighborhood of σ such that[4] $\bar{M} \subset N$. According to the preceding lemma there exists a neighborhood Y of σ such that $\bar{Y} \subset X$ and such that $A(X)$ is deformed on $\bar{Y}$ under $\Delta(t)$. It follows from the definition of $\Delta(t)$ that for t sufficiently large the final image U of $A(X)$ under $\Delta(t)$ will lie on $\bar{M} + \bar{Y}_c$. If the deformation $D(\tau)$ now be applied to U, the ordinary points of U at which $f = c$ will thereby be deformed into points below c. If this deformation be terminated at a sufficiently small value of $\tau > 0$, $\bar{M}$ and $\bar{Y}$ will not thereby be deformed off N and X respectively. Thus the resultant deformation will deform $A(X)$ on X onto $N + X_c$.

We shall prove the following lemma.

Lemma 4.3. *There exists a Δ-contractible neighborhood Γ of σ such that the Betti numbers of Γ_c are finite.*

Let Ω be an arbitrary neighborhood of σ. If e is a sufficiently small positive constant and N a sufficiently small neighborhood of σ on which $f > c - e$, the following statements will be true. From each point of N_c an orthogonal trajectory will lead on R to a point Q on $f = c - e$. The set S of these points Q will have a closure which lies on R and which consists of ordinary points of f on R. There will exist[5] a finite complex K on $f = c - e$ covering S. If K forms a sufficiently small neighborhood of S on $f = c - e$ and K' is the set of points on orthogonal trajectories issuing from K on which $c - e < f < c$, $\bar{N} + \bar{K}'$ will lie on $A(\Omega)$.

We set $\Gamma = N + K'$, and observe that Γ is Δ-contractible on Ω. We shall show that Γ_c has finite Betti numbers.

[4] The closure of a point set E will be denoted by $\bar{E}$.

[5] The methods of Cairns [3] suffice to prove this statement.

To that end we note that $\Gamma_c = K'$. With the aid of the orthogonal trajectories we see that the Betti numbers of K' are those of K, and accordingly finite.

The proof of the lemma is complete.

We introduce the following definition.

An ordered pair of neighborhoods VW of σ will be termed α-admissible if $V \subset \Gamma$ and W is Δ-contractible on V. We abbreviate the phrase 'corresponding to any α-admissible pair of neighborhoods VW' by the expression α-adm VW.

5. **Existence of maximal sets** α-adm VW. In this section we shall prove the following theorem.

THEOREM 5.1. *There exist maximal sets of spannable or critical k-cycles α-adm VW. The number of cycles in such sets is less than a finite constant independent of pairs VW which are α-admissible.*

Before coming to the proof of the theorem we define certain subclasses of spannable cycles.

A spannable $(k - 1)$-cycle u_{k-1}, α-adm VW, will be called *linkable* α-adm VW, if bounding on Γ_c. If u_{k-1} is linkable, there exists a chain λ_k'' on Γ_c such that

$$\lambda_k'' \rightarrow -u_{k-1} \qquad (\text{on } \Gamma_c)\,.$$

By virtue of the definition of a spannable $(k - 1)$-cycle, there also exists a chain λ_k' on W such that

$$\lambda_k' \rightarrow u_{k-1} \qquad (\text{on } W)\,. \tag{5.1}$$

We set

$$\lambda_k' + \lambda_k'' = \lambda_k\,, \tag{5.2}$$

and term λ_k a k-cycle *linking* u_{k-1}, α-adm VW. We shall say that λ_k belongs to σ.

A spannable k-cycle α-adm VW which is independent on Γ_c will be termed a *newly-bounding k-cycle* α-adm VW.

We begin the proof of Theorem 5.1 with the following lemma.[6]

LEMMA 5.1. *If Λ is a Δ-contractible neighborhood of σ and $V \subset \Lambda$, there exists no homology of the form*

$$m\lambda_k + nc_k + w_k \sim 0 \qquad (\text{on } \Lambda)\,, \tag{5.3}$$

where λ_k and c_k are respectively linking and critical k-cycles α-adm VW *and w_k is a k-cycle below c, unless $m \equiv n \equiv 0$, mod π.*

Suppose there were an homology of the form (5.3). We shall prove that $m \equiv n \equiv 0, \text{mod } \pi$.

The deformation $\Delta(t, \tau)$ applied for a sufficiently large time t and sufficiently small time τ will not only deform Λ onto $V + R_c$ but also deform W only on V.

[6] It follows from this lemma that there exists no relation of the form $m\lambda_k + nc_k + w_k = 0$ (on Λ) unless $m \equiv n \equiv 0$.

Let λ_k^*, c_k^* and w_k^* be, respectively, the final images of λ_k, c_k and w_k under $\Delta(t, \tau)$. We have the homologies

$$
\begin{aligned}
&\lambda_k - \lambda_k^* \sim 0 &&\text{(on } V + R_c),\\
(5.4)\qquad &c_k - c_k^* \sim 0 &&\text{(on } V),\\
&w_k - w_k^* \sim 0 &&\text{(on } R_c).
\end{aligned}
$$

Moreover, under $\Delta(t, \tau)$, (5.3) yields the homology

$$(5.5)\qquad m\lambda_k^* + nc_k^* + w_k^* \sim 0 \qquad \text{(on } V + R_c).$$

Substituting the homologies (5.4) in (5.5), we find that

$$(5.6)\qquad m\lambda_k + nc_k + w_k \sim 0 \qquad \text{(on } V + R_c).$$

We can write (5.6) in the form

$$z'_{k+1} + z''_{k+1} \to m\lambda_k + nc_k + w_k \qquad \text{(on } V + R_c),$$

where z'_{k+1} is a chain on V and z''_{k+1} is a chain on R_c. Let z'_k and z''_k be respectively the boundaries of z'_{k+1} and z''_{k+1}. Let u_{k-1} be the linkable $(k-1)$-cycle α-adm VW linked by λ_k. Recalling (5.2) and using the preceding bounding relations, we see that

$$(5.7)\qquad z'_k + z''_k \equiv m\lambda'_k + m\lambda''_k + nc_k + w_k \qquad (\text{mod } \pi).$$

We find that the chain

$$m\lambda'_k + nc_k - z'_k \qquad \text{(on } V)$$

reduced mod π, lies on R_c, since the remaining chains in (5.7) lie on R_c. From (5.1) we see that

$$(5.8)\qquad m\lambda'_k + nc_k - z'_k \to mu_{k-1} \qquad \text{(on } V).$$

Proof that $m \equiv 0$. If $m \not\equiv 0$ in (5.8), the spannable cycle mu_{k-1} bounds on V below c. This is contrary to the nature of a spannable cycle. Hence $m \equiv 0$.

Proof that $n \equiv 0$. Since $m \equiv 0$, it follows from (5.7) that the chain $z''_k - w_k$, reduced mod π, lies on V. Observe that $z'_k \sim 0$ on V. Hence we see from (5.7) that

$$(5.9)\qquad nc_k \sim z''_k - w_k \qquad \text{(on } V).$$

But $z''_k - w_k$ is a cycle below c. Hence (5.9) is contrary to the nature of the critical cycle c_k unless $n \equiv 0$. Thus $n \equiv 0$.

The proof of the lemma is complete.

Proof of Theorem 5.1. Since Γ is Δ-contractible, there will exist a neighborhood Λ of σ such that $\bar{\Lambda}$ is a finite complex, which covers Γ, and approximates Γ so closely as to be Δ-contractible. We make such a choice of Λ.

(a) *Critical k-cycles.* Let $(c)_k$ be a finite set of critical k-cycles α-adm VW, every proper linear combination c_k of whose cycles is of the same type. It

follows from Lemma 5.1 that $c_k \not\approx 0$ on Λ. Hence the number of cycles in $(c)_k$ is at most the k-th Betti number R_k of Λ. But R_k is finite. It follows that there exists a maximal set of critical k-cycles α-adm VW, and the number of cycles in such a set is at most R_k, a number which is independent of α-admissible pairs VW.

(b) *Linkable* $(k - 1)$*-cycles.* Let $(u)_{k-1}$ be a finite set of linkable $(k - 1)$-cycles α-adm VW, every proper linear combination u_{k-1} of whose cycles is of the same type. Let $(\lambda)_k$ be a set of k-cycles which link the respective $(k - 1)$-cycles of $(u)_{k-1}$. Let u_{k-1} be a proper linear combination of the cycles of $(u)_{k-1}$, and λ_k the corresponding linear combination of the cycles of $(\lambda)_k$. Recall that λ_k is a linking k-cycle α-adm VW. It follows from Lemma 5.1 that $\lambda_k \not\approx 0$ on Λ. Thus the number of cycles in $(\lambda)_k$, and hence in $(u)_{k-1}$, is at most the preceding number R_k. Hence there exists a maximal set of linkable $(k - 1)$-cycles α-adm VW and the number of cycles in such a set is bounded by R_k.

(c) *Newly-bounding k-cycles.* Let $(\nu)_k$ be a finite set of newly-bounding k-cycles α-adm VW, every proper linear combination of whose cycles is of the same type. It follows from the definitions of such cycles that they lie on Γ_c, and are independent on Γ_c. Hence the number of cycles of $(\nu)_k$ is at most the k-th Betti number P_k of Γ_c. It follows from Lemma 4.3 that P_k is finite. Hence there exists a maximal set of newly-bounding k-cycles α-adm VW, and the number of cycles in such a set is at most P_k, a number which is independent of α-admissible pairs VW.

Theorem 5.1 follows from (a) in so far as it refers to critical cycles. With respect to spannable cycles, observe that a maximal set of spannable k-cycles α-adm VW will be afforded by the sum of maximal sets of linkable and newly-bounding k-cycles α-adm VW, as follows from the definitions of the cycles involved. The theorem accordingly holds for spannable cycles as well.

6. **Proof of the Neighborhood Theorem.** In this section we shall establish the Neighborhood Theorem of §3. Our first lemma is the following.

LEMMA 6.1. *Corresponding to an arbitrary neighborhood V of σ on Γ, there exists a neighborhood $B(V)$ of σ on $A(V)$ such that if*

$$W \subset B(V),$$

a maximal set of critical or spannable k-cycles α-adm VW *is also a maximal set* α-adm $V\,B(V)$.

The lemma will be proved for the case of critical cycles. The proof for the case of spannable cycles is similar.

Let W and W' be any two neighborhoods of σ for which $W \subset W' \subset A(V)$. Let $(c)_k$ be a maximal set of critical k-cycles α-adm VW. It follows at once from the definitions of critical k-cycles and of maximal sets that $(c)_k$ is a subset of a maximal set of critical k-cycles α-adm VW'. But the number of cycles in a maximal set of critical k-cycles is finite. Hence if W' is taken as a sufficiently small neighborhood $B(V)$ of σ, the number of cycles in $(c)_k$ will be independent

of W, provided $W \subset B(V)$, and $(c)_k$ will be a maximal set of critical k-cycles α-adm $V\,B(V)$. The proof of the lemma is complete.

We continue with the following lemma.

LEMMA 6.2. *There exists a fixed neighborhood H of σ on Γ such that if*

$$V \subset H, \qquad W \subset B(H)\cdot B(V), \tag{6.1}$$

a maximal set of critical or spannable k-cycles α-adm HW *is a maximal set* α-adm VW.

We consider the case of critical k-cycles. Suppose V and V' are any two neighborhoods of σ for which $V \subset V' \subset \Gamma$, and W is any neighborhood of σ on $B(V')\cdot B(V)$. The pairs VW and $V'W$ are then α-admissible. Theorem 5.1 applies, and maximal sets of critical cycles exist. Moreover, it follows from the definition of critical cycles that a maximal set $(c)_k$ of critical k-cycles α-adm $V'W$ is a subset of a maximal set of critical k-cycles α-adm VW. Recall that the numbers of cycles in such sets α-adm $V'W$ are bounded with respect to all pairs $V'W$ which are α-admissible, and with the above choice of W depend only on V', by virtue of Lemma 6.1. If then V' is taken as a sufficiently small neighborhood H of σ on Γ, the set $(c)_k$ α-adm $V'W$ will be a maximal set α-adm VW, and the lemma holds as stated.

We introduce the following definition.

A pair of neighborhoods VW of σ will be termed β-admissible if (6.1) *holds. The phrase* β-adm VW *will have the obvious meaning.*

We come to the following theorem.

THEOREM 6.1. *Corresponding to any two β-admissible pairs of neighborhoods VW and $V'W'$ of σ there exist on any arbitrarily small neighborhood of σ common maximal sets of spannable and critical k-cycles.*

Let N be any neighborhood of σ such that $N \subset W\cdot W'$. Let $(c)_k$ be a maximal set of critical k-cycles β-adm HN. Such a set exists by virtue of Theorem 5.1. It follows from Lemma 6.2 that $(c)_k$ is a maximal set of critical k-cycles both β-adm VN and β-adm $V'N$. It then follows from Lemma 6.1 that $(c)_k$ is a maximal set of critical k-cycles both β-adm VW and β-adm $V'W'$. Hence the theorem holds for critical cycles.

The proof for the case of spannable cycles is similar.

Proof of the Neighborhood Theorem (§3). If we choose the neighborhoods N^* and $M(X)$ of the Neighborhood Theorem respectively as the neighborhoods H and $B(H)\cdot B(X)$ defined above, we see that the Neighborhood Theorem is an immediate consequence of Theorem 6.1.

We introduce the following definition.

An ordered pair of neighborhoods VW of σ will be termed admissible if they satisfy the conditions

$$V \subset M(N^*), \qquad W \subset M(V),$$

where N^* and $M(X)$ are the neighborhoods of the Neighborhood Theorem.

The phrase adm VW will have the obvious meaning.

It follows at once from the Neighborhood Theorem that the total number, say m_k, of cycles in maximal sets of spannable $(k-1)$-cycles and critical k-cycles adm VW is independent of the choice of admissible pairs of neighborhoods VW. We define the number m_k to be the k-th *type number* of the critical set σ.

The type numbers of σ depend on the definition of f only in an arbitrarily small neighborhood of σ. It is by virtue of the Neighborhood Theorem that the type numbers exist and are finite.

7. **Group aspects in the small.** In this section we shall discuss the group theory aspects of cycles neighboring σ.

We shall deal with a quotient[7] $G = A/B$ of groups of cycles A and B in which the operation is addition mod π. This will be understood throughout and need not be repeated. In the group G the operation will again be addition. The elements of G will be classes of cycles, two cycles belonging to the same class if their difference mod π belongs to the class which is the zero-element of G. Any class which is not the zero-element of G will be called a *proper* class. A cycle belonging to a class C will be termed a *representative* of C. We shall deal with groups which possess a finite number of generators.

Let VW be an arbitrary pair of neighborhoods of σ of which $W \subset V$. We define the following groups corresponding to VW.

Let A be the group of k-cycles on W below c, ~ 0 on W, and B the subgroup of cycles dependent on V below c. We term A/B the k-th *spannable group corresponding to* VW.

Similarly let A be the group of k-cycles on W, and B the subgroup of cycles $\approx$ on V to cycles below c. We term A/B the k-th *critical group corresponding to* VW.

We have the following theorem.

THEOREM 7.1. *A cycle is a spannable (critical) k-cycle corresponding to VW if and only if it is a representative of a proper class of the k-th spannable (critical) group corresponding to VW.*

Let VW be an α-admissible pair (§4) of neighborhoods of σ. We define the k-th *linkable group* α-adm VW to be the group of those classes of the k-th spannable group α-adm VW whose representative cycles are dependent on Γ_c.

The quotient of the k-th spannable group by the k-th linkable group α-adm VW will be termed the k-th *newly-bounding group* α-adm VW. We can and shall regard the elements of the latter group as classes of cycles.

We have the following theorem.

THEOREM 7.2. *A cycle is a newly-bounding k-cycle* α-adm VW *if and only if it is a representative of a proper class of the k-th newly-bounding group* α-adm VW.

[7] For additive groups, quotients such as A/B are frequently written as A mod B.

II. The fundamental relations

8. **The function and the region.** We turn now to the problem of determining the fundamental relations between the type numbers of the critical sets of f and the Betti numbers of the domain of definition of f.

Let Σ be a limited, open, n-dimensional subregion of the region R of §2, whose closure $\bar{\Sigma}$ lies on R, and whose boundary B is a closed point set consisting of a finite number of connected, regular, non-intersecting $(n - 1)$-spreads[8] of class C^3. Let f_ν denote the directional derivative of f on the normal to B in the sense that leads from points on Σ to points not on Σ.

A function f will be termed A-*admissible on* Σ if it satisfies the following conditions.

A I. *The function f shall be of class C^1L on an open region containing $\bar{\Sigma}$ and shall have only a finite number of critical values on Σ.*

A II. *The function f shall be of class C^2 neighboring B. The directional derivative f_ν of f shall be positive on B.*

We assume that the function f is A-admissible on Σ. As in [9], we can alter the definition of f neighboring B so that the resulting function, which we will again call f, is A-admissible on Σ, but in addition is constant on B, its value on B being greater than at any point of Σ. This alteration can be made without introducing any new critical points. We assume that f has been altered in this way. From this point on we consider f only on $\bar{\Sigma}$.

If a and b are two ordinary values of f, with no critical values between them, the domains $f \leqq a$ and $f \leqq b$ are homeomorphic, [9]. When there are critical values between a and b this will not in general be so. We are concerned in what follows with the topological differences between the domains $f \leqq a$ and $f \leqq b$, and the manner in which these differences depend on the critical points of f.

We recall the definitions of critical sets and their neighborhoods made in §2. Let D be a closed subdomain of Σ on whose boundary f has no critical points. The set ω of all critical points of f on D at which $f = c$ will be termed a *complete* critical set *on D*. In general the points of ω may be grouped into a finite number of disjoint critical sets in several ways, but it follows from the definition of critical sets that it is not possible to group the points of ω into an infinite ensemble of disjoint critical sets. Suppose ω is the sum of disjoint critical sets $\omega_1, \cdots, \omega_m$. It follows from the definition of critical sets that the admissible pairs of neighborhoods used in defining the type numbers of the several sets ω_i may be chosen so small that pairs belonging to distinct sets ω_i are disjoint. Hence the k-th type number of ω equals the sum of the k-th type numbers of the several sets ω_i, and is thus independent of the way in which ω is broken up into a sum of a finite number of disjoint critical sets.

[8] An $(n - 1)$-spread is said to be regular and of class C^r $(r > 0)$ if in the neighborhood of each of its points one of its coördinates can be represented as a function of class C^r of the remaining coördinates.

9. **Classification of cycles.** Let c be a critical value of f, and $b > c$ an ordinary value, such that between c and b there is no critical value of f. In this section we shall give a maximal set of k-cycles on $f < b$, independent on $f < b$, in terms of k-cycles belonging to the complete critical set ω of f on $f = c$.

Let σ be any critical set of f on which $f = c$. Linkable, linking and newly-bounding k-cycles adm VW are formally defined as in §5, the domain Γ_c being replaced by the domain $f < c$. A linking k-cycle λ_k adm VW which links a linkable $(k - 1)$-cycle u_{k-1} adm VW may be represented in the form

$$\lambda_k = \lambda_k' + \lambda_k'', \tag{9.1}$$

where λ_k' is a chain on W and λ_k'' a chain on R_c, the boundaries of λ_k' and λ_k'' being u_{k-1} and $-u_{k-1}$ respectively.

We come to four lemmas on linking and critical cycles. We begin with the following.

Lemma 9.1. *Let $(u)_{k-1}$ be a set of linkable $(k - 1)$-cycles* adm VW, *and $(\lambda)_k$ a set of k-cycles linking the respective $(k - 1)$-cycles of the set $(u)_{k-1}$* adm VW. *A necessary and sufficient condition that $(u)_{k-1}$ be a maximal set of linkable $(k - 1)$-cycles* adm VW *is that $(\lambda)_k$ be a maximal set of linking k-cycles* adm VW.

The proof of this lemma is essentially the same as that of Lemma 6.1, [8]. One need only replace the symbol $\sim$ by $\not\approx$, the word "sum" by the phrase "proper linear combination," together with certain changes of sign depending upon the substitution of mod π for mod 2.

Our next lemma follows.

Lemma 9.2. *If VW is an admissible pair of neighborhoods of σ, any k-cycle on V is $\approx$ on N^* to a linear combination of critical k-cycles* adm VW *and cycles below c.*

It follows at once from the definitions that a k-cycle on V is $\approx$ on N^* to a sum of critical k-cycles corresponding to N^* $M(N^*)$ and cycles below c. But it follows from the Neighborhood Theorem that a maximal set of critical k-cycles corresponding to N^* $M(N^*)$ may be taken as the cycles of a maximal set of critical k-cycles adm VW.

The proof of the lemma is complete.

Lemma 9.3. *Let $(\lambda)_k$ be a maximal set of k-cycles linking* adm VW. *Then any k-cycle on $W + R_c$ is $\approx$ on $N^* + R_c$ to a linear combination of cycles of $(\lambda)_k$, critical k-cycles* adm VW *and k-cycles below c.*

Let $(u)_{k-1}$ be the set of $(k - 1)$-cycles linked respectively by the cycles of the set $(\lambda)_k$. Let z_k be an arbitrary k-cycle on $W + R_c$. If sufficiently finely divided z_k can be represented in the form

$$z_k = z_k' + z_k'', \tag{9.2}$$

where z_k' is a chain on W and z_k'' a chain on R_c. Suppose z_{k-1} is the boundary of z_k'. The cycle z_{k-1} is below c. It accordingly satisfies an homology

$$z_{k-1} \approx r u_{k-1} \qquad \text{(on } V \text{ below } c\text{)}, \tag{9.3}$$

where u_{k-1} is a proper linear combination of cycles of $(u)_{k-1}$, and $r = 1$ or 0.

Let u_{k-1} be a proper linear combination of cycles of $(u)_{k-1}$, and λ_k the corresponding linear combination of the cycles of $(\lambda)_k$. By virtue of (9.3) there exists a chain w_k on V below c and an integer m such that

$$w_k \to mz_{k-1} - mru_{k-1} \qquad (m \not\equiv 0).$$

Upon using (9.1) and (9.2) we find that

$$(9.4) \qquad mz_k - mr\lambda_k \equiv (mz_k' - mr\lambda_k' - w_k) + (mz_k'' - mr\lambda_k'' + w_k).$$

The first parenthesis contains a k-cycle on V and the second a k-cycle below c. But k-cycles on V are $\approx$ on N^* to sums of critical k-cycles adm VW and cycles below c. Since the modulus π is prime, the congruence (9.4) yields the lemma.[9]

We return to the complete critical set ω of f on $f \leqq b$, on which $f = c$. We define a new set of cycles.[10] A k-cycle below c independent below c, and independent below c of the spannable k-cycles adm VW is termed an *invariant k-cycle* adm VW.

We come to a basic theorem. [8, Theorem 6.1.]

THEOREM 9.1. *A maximal set of k-cycles on $f < b$, independent on $f < b$, is afforded by maximal sets of critical, linking and invariant k-cycles corresponding to an admissible pair of neighborhoods VW of the complete critical set ω on which $f = c$.*

We shall prove the theorem by proving statements (a) and (b). Statement (a) follows.

(a) *Any k-cycle z_k on $f < b$ is $\approx$ on $f < b$ to a linear combination of the k-cycles of the maximal sets* adm VW *of the theorem.*

By means of the deformation $D(t)$ of §4, the domain $f < b$ can be deformed on itself onto the domain $W + R_c$. Hence we lose no generality if we suppose that z_k lies on $W + R_c$. It follows from the definition of invariant cycles that cycles below c are $\approx$ on $f < b$ to a linear combination of invariant cycles. Statement (a) follows at once from Lemma 9.3.

(b) *The cycles of the maximal sets of the theorem are independent on $f < b$.*

Suppose that there existed an homology of the form

$$(9.5) \qquad m\lambda_k + nc_k + ri_k \sim 0 \qquad (\text{on } f < b),$$

where $m, n, r = 1$ or 0, and where λ_k, c_k, and i_k are respectively linking, critical and invariant k-cycles adm VW.

It follows, as in the proof of Lemma 5.1, that the homology (9.5) may be taken on $V + R_c$. Continuing as in the proof of Lemma 5.1, we infer that $m = n = 0$. Hence (9.5) implies the homology

$$(9.6) \qquad ri_k \sim 0 \qquad (\text{on } f < b).$$

[9] Were π not a prime, the congruence derived from (9.4) by replacing multiples of the parentheses on the right by the desired cycles might contain a leading coefficient on the right of the form $mn \equiv 0$, and the proof would fail.

[10] In what follows it is convenient to suppose that all neighborhoods of ω which are admitted lie on $f < b$.

Since i_k is on R_c, we can deform $f < b$ on itself onto $W + R_c$, keeping i_k fixed. Hence (9.6) implies the homology $ri_k \sim 0$ (on $W + R_c$). Let z_{k+1} be the chain on $W + R_c$ bounded by ri_k. We can write

$$z'_{k+1} + z''_{k+1} \to ri_k, \tag{9.7}$$

where z'_{k+1} is a chain on W and z''_{k+1} a chain on R_c. Let z'_k and z''_k be, respectively, the boundaries of z'_{k+1} and z''_{k+1}. From (9.7) we see that

$$ri_k - z'_k - z''_k \equiv 0.$$

Since i_k and z''_k are on R_c, the cycle z'_k is on R_c. And $z''_k \sim 0$ on R_c. Hence

$$ri_k \sim z'_k \qquad \text{(on } R_c). \tag{9.8}$$

If $z'_k = 0$ on R_c, r must be zero in (9.8), for invariant cycles are independent below c. If $z'_k \not\approx 0$ on R_c, z'_k is a spannable k-cycle adm VW. We again infer that $r = 0$, since invariant k-cycles are independent below c of spannable k-cycles adm VW.

Thus in (9.5), $m = n = r = 0$, and the proof of (b) is complete. The theorem follows directly. Cf. note 6 following Lemma 5.1.

10. **The relations.** We come to the relations between the type numbers of the critical sets of f and the Betti numbers of Σ.

Recall that ω is the complete critical set of f on Σ on which $f = c$. Let a and b ($a < b$) be ordinary values of f between which c is the only critical value. Relative to the critical value c and the constants a and b, a *new k-cycle* shall mean a k-cycle on $f < b$, not $\approx$ on $f < b$ to k-cycles on $f < a$. Relative to the critical value c and the constants a and b, a *newly-bounding k-cycle* shall mean a k-cycle on $f < a$, independent on $f < a$, but bounding on $f < b$.

Since the Betti numbers of the domains $f < a$ and $f < b$ are finite, it follows that maximal sets of new k-cycles and newly-bounding $(k - 1)$-cycles exist. Moreover, the numbers of cycles in such maximal sets are independent of ordinary values a and b ($a < b$), between which c is the only critical value. Let m_k^+, m_k^- be respectively the numbers of cycles in maximal sets of new k-cycles and newly-bounding $(k - 1)$-cycles relative to the critical value c.

Recall that a newly-bounding k-cycle adm VW is a spannable k-cycle adm VW which is independent below c. It follows from this and from the definition of invariant $(k - 1)$-cycles adm VW that a maximal set of $(k - 1)$-cycles independent below c consists of maximal sets of invariant and newly-bounding $(k - 1)$-cycles adm VW. Of these cycles the invariant $(k - 1)$-cycles adm VW remain independent on $f < b$, according to Theorem 9.1. Hence m_k^- equals the number of newly-bounding $(k - 1)$-cycles adm VW in a maximal set of such cycles. It also follows from Theorem 9.1 that m_k^+ is the number of critical and linking k-cycles in maximal sets of such cycles adm VW. It follows from Lemma 9.1 that the numbers of cycles in maximal sets of linking k-cycles and linkable $(k - 1)$-cycles adm VW are the same. Observe finally that maximal

sets of linkable and newly-bounding $(k-1)$-cycles together form a maximal set of spannable $(k-1)$-cycles.

Combining the above statements with the definition in §6 of the type numbers m_k of ω, we see that $m_k = m_k^+ + m_k^-$. We are thus led to the following theorem.

THEOREM 10.1. *Let a and b, $a < b$, be two ordinary values of f between which c is the only critical value. Let ΔR_k denote the k-th Betti number of the domain $f < b$ minus that of the domain $f < a$. Let m_k be the k-th type number of the complete critical set ω on which $f = c$. Finally, let m_k^+ and m_k^- be, respectively, the numbers of cycles in maximal sets of new k-cycles and newly-bounding $(k-1)$-cycles relative to the critical value c. Then*

$$\Delta R_k = m_k^+ - m_{k+1}^-, \qquad m_k = m_k^+ + m_k^- \qquad (k = 0, 1, \cdots, n),$$

where $m_0^- = m_{n+1}^- = m_n^+ = 0$.

The following theorem, which is an immediate consequence of Theorem 10.1, affirms the validity in the present case of the earlier relations of Morse [8, Theorem 1.1]. In it f is a function which is A-admissible on Σ.

THEOREM 10.2. *Between the Betti numbers R_i of Σ and the sums M_i ($i = 0, 1, \cdots, n$) of the i-th type numbers of the critical sets of f on Σ, the following relations hold:*

$$R_0 \leqq M_0,$$

$$R_0 - R_1 \geqq M_0 - M_1,$$

$$R_0 - R_1 + R_2 \leqq M_0 - M_1 + M_2,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$R_0 - R_1 + R_2 - \cdots + (-1)^n R_n = M_0 - M_1 + M_2 - \cdots + (-1)^n M_n.$$

11. **Group aspects in the large.** We shall examine the results of §9 from the standpoint of groups.

Let VW be an admissible pair of neighborhoods of the complete critical set ω of §9. We make the following definitions.

Let A be the group of k-cycles on $W + R_c$ and B the subgroup of cycles $\approx$ on $N^* + R_c$ to cycles on V plus cycles below c. We term A/B the k-th *linking group* adm VW.

Similarly, let A be the group of k-cycles below c and B the subgroup of these cycles ≈ 0 on $f < b$. Relative to the critical value c we term A/B the k-th *invariant group*.

Finally, let A be the group of k-cycles on $f < a$ [$f < b$], and B the subgroup of cycles ≈ 0 on $f < a$ [$f < b$]. One terms A/B the k-th *Betti group* of $f < a$ [$f < b$].

We have the following theorems.

THEOREM 11.1. *A k-cycle on $W + R_c$ is a linking k-cycle* adm VW *if and only if it is a representative of a proper class of the k-th linking group.*

THEOREM 11.2. *A k-cycle below c is an invariant k-cycle* adm VW *if and only if it is a representative of a proper class of the k-th invariant group.*

Theorem 9.1 takes the following form.

THEOREM 11.3. *The k-th Betti group of $f < b$ is isomorphic with the direct sum of the k-th linking group* adm VW, *the k-th critical group* adm VW, *and the k-th invariant group.*

III. General boundary conditions

12. Admissible functions. We come to a generalization of the conditions A of §8. Let the region Σ and its boundary B be defined as in §8.

Let P be any point on B. It follows from the regularity of B that B can be represented neighboring P in the form

$$x_i = g_i(u^1, \cdots, u^{n-1}) \qquad (i = 1, \cdots, n),$$

where the functions g_i are of class C^3 neighboring a point (u_0) that determines P, and where the matrix of the first partial derivatives of the functions g_i is of rank $n - 1$ at (u_0). Parameters (u) in such a local representation of B will be termed *admissible.*

The function f^0 defined by f on B will be termed the boundary function defined by f on B.

We shall represent f^0 in terms of parameters of B, using, however, only those parameters of B which we have termed admissible.

Let (u) be a set of parameters admissibly representing B neighboring a point P on B, and let $\psi(u)$ be the value of f^0 in terms of these parameters (u). The critical points of f^0 neighboring P are defined by the critical points of $\psi(u)$. The critical points of f^0 are clearly independent as points on B of the parameters (u) which are used locally to represent B. The critical sets of f^0 on B are defined in the same manner as the critical sets of f on R.

We shall subsequently refer to neighborhoods of the critical sets of f^0. We understand that these neighborhoods are open in B, in the point set sense.

A function $f(x)$ will be termed B-*admissible on* Σ if it satisfies the following conditions.

B I. *The function f shall satisfy condition* A I.

B II. *The function f shall be of class C^3 neighboring B and shall have no critical points on B. The boundary function f^0 shall have a finite number of critical values.*

By making use of the trajectories orthogonal to the manifolds $f^0 =$ constant, we can define a deformation on B similar to the deformation $\Delta(t)$ of §3. Δ-contractible neighborhoods on B of a critical set σ^0 of f^0 can then be defined essentially as in §3 and a theorem established similar to the Neighborhood Theorem in §3. Admissible pairs V^0W^0 of neighborhoods of σ^0 on B and the type numbers of σ^0 can be defined essentially as before. *Complete* critical sets of f^0 on closed domains D of B are defined as were complete critical sets of f on closed domains D of Σ.

Let β be a connected boundary spread of B. We admit the possibility that the boundary function f^0 be identically a constant on β. Suppose then that

$f^0 \equiv c$ on β. The spread β is a critical set of f^0. A sufficiently small neighborhood of β on B will be identical with β. Neighboring β on B there are no points below c and hence no spannable cycles belonging to β. Each cycle on β which is non-bounding on β will be a critical cycle belonging to β. The number of critical k-cycles in a maximal set of such cycles belonging to β will equal the k-th Betti number of β.

Recall that f_ν is the directional derivative of f on the normal to B in the sense that leads from points on Σ to points not on Σ. The set of points on B at which $f_\nu \leqq 0$ will be termed the *negative boundary* of Σ.

13. **The principal theorem.** The following theorem gives the main results of this part of the paper. In it $f(x)$ is a function which is B-admissible on Σ.

THEOREM 13.1. *Let M_i ($i = 0, 1, \cdots, n$) be the sum of the i-th type numbers of the complete critical sets of f on Σ and of the boundary function on the negative boundary of Σ, and let R_j ($j = 0, 1, \cdots, n$) be the j-th Betti number of $\bar{\Sigma}$. The numbers M_i and R_j satisfy the relations in Theorem* 10.2. Cf. Theorem 1.1, [8].

This theorem is obtained by applying the results of Theorem 10.2 to the function $F(x)$ described in the following lemma. In this lemma we let σ^0 denote an arbitrary complete critical set of f^0 on the negative boundary of $\bar{\Sigma}$.

FUNDAMENTAL LEMMA. *On a suitably chosen open region including $\bar{\Sigma}$ there exists a function $F(x)$ which is identical with $f(x)$ on $\bar{\Sigma}$ except neighboring B, and which is* A-*admissible on Σ.*

The critical points of $F(x)$, other than those of $f(x)$, may be grouped into critical sets $\sigma(\sigma^0)$ which correspond in a one-to-one manner to the critical sets σ^0, and possess type numbers $m_0, \cdots, m_{n-1}$ which equal the corresponding type numbers of $\sigma(\sigma^0)$, while $m_n = 0$.

The remaining sections of the paper will be occupied with the proof of this lemma. In §14 we shall modify f at the points on Σ neighboring the critical sets of f^0 on the negative boundary of Σ in such a way that on B neighboring these critical sets the new function will have a constant negative normal directional derivative. In §15 we shall extend the definition of the new function to points neighboring B outside Σ. The function H thus obtained will have a positive normal directional derivative on a spread B_1 neighboring B. The function H is of the type already studied in §§8–10. Its critical sets are determined by those of f and of f^0, as we shall see in §16.

14. **Modification of f within Σ.** In this section we shall make use of special coördinates neighboring B. Let (x^0) be the rectangular coördinates (x) of a point P on B. It can be shown that the normals to B neighboring P form a field neighboring P. Let (u) be a set of admissible parameters in terms of which B is locally representable neighboring P. Let (u_0) be the parameters of P. On each normal to B neighboring P let s be the arc-length measured positively in the sense of the outer normal, with $s = 0$ on B. There exists a positive constant η and on B a neighborhood M of (u_0) such that on the normals to B

at points of M the points at which $|s| \leqq \eta$ cover a neighborhood M^* of (x^0) in a one-to-one manner. Sets (u, s) for which (u) is on M and $|s| \leqq \eta$ may be considered as new coördinates for the points of M^*. The transformation from the coördinates (x) to the coördinates (u, s) will be of class C^2 and will have a non-vanishing jacobian. We shall call such coördinates (u, s) *admissible* neighboring P. We can suppose the constant η is so small that the points on the normals to B at which $|s| \leqq \eta$ cover a neighborhood of B in a one-to-one manner include no critical points of f, and include only points at which f is of class C^3.

It is convenient at this point to define several functions.

The invariant function φ *on* B. In terms of local coördinates (u) on B let

$$d\sigma^2 = g_{ij}\, du^i\, du^j \qquad (i, j = 1, \cdots, n-1)$$

be the differential quadratic form defining the metric on B. Let $\psi(u)$ be the local representation of the boundary function f^0 and let $\psi_i = \dfrac{\partial\psi}{\partial u^i}$. The function[11]

$$\varphi = g^{ij}\psi_i\psi_j \qquad (i, j = 1, \cdots, n-1)$$

is an invariant function on B and assumes a proper minimum on each critical set σ^0 of the boundary function f^0.

The functions $l(z)$ *and* $\lambda(z)$. Let $l(z)$ be a function of class C^3 for all z such that

$$\begin{aligned} l(0) &= 0, & & \\ l(z) &\equiv 0, & & 4 \leqq z^2, \\ l'(0) &= 1, & & \\ l'(z) &\geqq 0, & & 0 < z^2 \leqq 1, \\ -\alpha \leqq l'(z) &\leqq 0, & & 1 < z^2 \leqq 4, \\ l''(0) &= 0, & & \end{aligned}$$

where α is an arbitrary positive constant. Let $\lambda(z)$ be a function of class C^3 for all z such that

$$\begin{aligned} \lambda(z) &\equiv 1, & & 0 \leqq z^2 \leqq 1, \\ 0 \leqq \lambda(z) &\leqq 1, & & 1 < z^2 < 4, \\ \lambda(z) &\equiv 0, & & 4 \leqq z^2. \end{aligned}$$

Functions $l(z)$ and $\lambda(z)$ can easily be constructed.

The functions $g(x_1, \cdots, x_n)$ *and domains* U_r. Let σ^0 be a complete critical set of f^0 on the negative boundary of Σ. If r is a sufficiently small positive constant, the points of B connected to σ^0 at which the invariant function $\varphi < 3r$

[11] See Eisenhart, *Riemannian Geometry*, Princeton, 1926, p. 14.

form a neighborhood U_r^0 of σ^0. Let U_r be the domain of points neighboring B which are on the normals to B at points of U_r^0 and for which

$$0 \geqq s > -3r \qquad (3r < \eta).$$

We suppose the constant r is so small that the domains U_r corresponding to the complete critical sets σ^0 with different critical values are disjoint and that $f_\nu < 0$ on the closure of each of these domains. Such values of r we term *admissible.*

On each domain U_r we define a function $g(x)$. Let f_ν^0 be the value of f_ν on B and let $M < 0$ be a lower bound of f_ν^0 on B. Let P be a point of U_r on the normal to B at a point Q. Let (u, s) be admissible coördinates neighboring P. Let $\Phi(u, s)$ be the local representation of f neighboring P and $\Psi(u)$ the local representation of f_ν^0 neighboring Q. Then neighboring P at the point (x) determined by the coördinates (u, s), $g(x)$ shall have the representation

$$g(x) = \Phi(u, s) + rl\left(\frac{s}{r}\right)\lambda\left(\frac{\varphi(u)}{r}\right)[M - \Psi(u)] \qquad (-3r < s \leqq 0),$$

where $l(z)$ is defined as previously with α taken as r.

One readily verifies the fact that $g(x)$ is of class C^2. The directional derivative of g along the outer normal to B is

$$g_\nu = \Phi_\nu + l'\left(\frac{s}{r}\right)\lambda\left(\frac{\varphi}{r}\right)[M - \Psi]. \tag{14.1}$$

On U_r^0, $s = 0$, and for $\varphi \leqq r$, $\lambda\left(\dfrac{\varphi}{r}\right) \equiv 1$.

Hence at points of U_r^0 neighboring σ^0 for which $\varphi \leqq r$

$$g_\nu = M.$$

For r sufficiently small the corresponding function g has no critical points. For on U_r the function f_ν is bounded above by a negative constant, say $-m$. On the same domain the second term in the right member of (14.1) is bounded above by the positive number $-r^2M$. Thus if we choose an admissible value of r so small that $| r^2M | < m$, we shall have $g_\nu < 0$ on U_r. Hence g will have no critical points.

Finally, on the subdomain of U_r exterior to the domain

$$\varphi < 2r, \qquad -2r < s \leqq 0,$$

the function $g(x)$ is identical with $f(x)$.

The function $G(x_1, \cdots, x_n)$. On $\bar{\Sigma}$ we define a function $G(x)$ which is a modification of $f(x)$. At a point (x) of $\bar{\Sigma}$ not on any of the several domains U_r we set $G(x) = f(x)$. At a point (x) of a domain U_r we set $G(x) = g(x)$, where $g(x)$ is the function corresponding to U_r.

One readily verifies the statements in the following lemma.

LEMMA 14.1. *The function $G(x)$ is of class C^1L on Σ and is identical with $f(x)$ on B. In general $G(x)$ is identical with $f(x)$ except in neighborhoods of the complete critical sets σ^0 of f^0 on the domain $f_\nu^0 \leqq 0$. At points on B neighboring these critical sets σ^0, G_ν is the negative constant M.*

15. **Modification of f outside Σ.** Let s_1 and s_2 be two arbitrarily small positive constants with $0 < s_1 < s_2 < \eta$, where η is the constant used in §14. Let Σ_i $(i = 1, 2)$ denote the points of Σ together with the points neighboring B for which $0 \leqq s < s_i$. Let B_1 denote the boundary of Σ_1.

The function $h(x_1, \cdots, x_n)$. Let S denote the points neighboring B for which $0 \leqq s < s_2$. On S we shall define a function $h(x)$. Recall that M is a negative lower bound of f_ν^0 on B. Let G_ν^0 be the value on B of the directional derivative of $G(x)$ along the outer normal to B, where $G(x)$ is the function of §14. Let L be a positive constant larger than $-M$. Let P be a point on S on the normal to B at a point Q. Let (u, s) be admissible coördinates neighboring P. Let $\psi(u)$ be the local representation of f^0 neighboring Q and $\chi(u)$ the local representation of G_ν^0 neighboring Q. Then neighboring P at the point (x) determined by the coördinates (u, s) the function $h(x)$ shall have the representation

$$h(x) = f^0 + sG_\nu^0 + \frac{Ls^3}{3s_1^2} = \psi(u) + s\chi(u) + \frac{Ls^3}{3s_1^2} \qquad (0 \leqq s < s_2). \tag{15.1}$$

One sees that $h(x)$ is of class C^2.

The function $H(x)$. On Σ_2 we define a function $H(x)$. To that end we set

$$H(x) \equiv G(x) \qquad (\text{on } \Sigma),$$

$$H(x) \equiv h(x) \qquad (\text{on } S).$$

We shall prove the following lemma.

LEMMA 15.1. *The function $H(x)$ is* A*-admissible on Σ_1. The functions $H(x)$ and $f(x)$ are identical neighboring their respective critical sets on Σ.*

The function $H(x)$ is of class C^2 on Σ neighboring B and on S. One readily shows that $H(x)$ is of class C^1L neighboring B and hence of class C^1L on Σ_2. The normals to B are also normals to B_1. On B_1 the directional derivative H_ν along the outer normal to B_1 is positive. For from (15.1) we find that

$$H_\nu = G_\nu^0 + L \qquad (\text{on } B_1).$$

But since $G_\nu^0 \geqq M$, we see that

$$H_\nu \geqq M + L > 0 \qquad (\text{on } B_1).$$

The proof of the lemma is complete.

16. **The critical sets of H neighboring B.** It is clear that the only critical points of $H(x)$ other than those of $f(x)$ occur at points neighboring B for which $0 < s < s_2$. In this section we shall discuss the existence of such critical points.

Let σ^0 be a complete critical set of f^0 on the negative boundary of Σ. Let the constant r have the value used in defining the function $G(x)$ of §14. The points of B connected to σ^0 for which the invariant function $\varphi < r$ form a neighborhood U^* of σ^0. Let U denote the domain of points neighboring B which *project* orthogonally by means of the normals to B into U^* and for which $0 \leqq s < s_2$.

We now set

$$s^* = \sqrt{\frac{-M}{L}}\, s_1$$

and prove the following lemma.

LEMMA 16.1. *The points of the point set σ^* on U for which $s = s^*$ and which project orthogonally into the points of σ^0 on B are the only critical points of $H(x)$ on U.*

Let a neighborhood of each point on U be represented by admissible coördinates (u, s). The conditions which define the critical points (u, s) of H neighboring a point of U are, according to (15.1),

$$(16.1)' \qquad \frac{\partial H}{\partial u^i} = \frac{\partial f^0}{\partial u^i} + s\,\frac{\partial G^0_\nu}{\partial u^i} = 0 \qquad (i = 1, \cdots, n-1; \varphi < r),$$

$$(16.1)'' \qquad \frac{\partial H}{\partial s} = G^0_\nu + \frac{Ls^2}{s_1^2} = 0 \qquad (\varphi < r).$$

The points of σ^* satisfy $(16.1)'$ as follows from the fact that at each point of σ^0 on B

$$\frac{\partial f^0}{\partial u^i} = 0, \qquad G^0_\nu = M.$$

The second condition $(16.1)''$ bears upon s alone and is satisfied by $s = s^*$. Hence the points of σ^* are critical points of $H(x)$.

Moreover, the condition $(16.1)''$ is satisfied only by $s = s^*$. On the other hand, $(16.1)'$ is satisfied only when (u) represents a point of σ^0, because for $0 < \varphi < r$

$$\frac{\partial G^0_\nu}{\partial u^i} = \frac{\partial M}{\partial u^i} = 0,$$

and for at least one value of i

$$0 \neq \frac{\partial f^0}{\partial u^i} = \frac{\partial H}{\partial u^i} \qquad (0 < \varphi < r).$$

The proof of Lemma 16.1 is complete.

The point set σ^* of Lemma 16.1 will be termed the critical set of $H(x)$ *corresponding* to the critical set σ^0 of f^0.

We continue with the following lemma.

LEMMA 16.2. *For s_2 sufficiently small the function $H(x)$ has no critical points*

for which $0 \leqq s < s_2$ *except the points of the critical sets* σ^* *corresponding to the complete critical sets* σ^0 *of* f^0 *on the negative boundary of* Σ.

Let B' denote the complement on B of the sum of the preceding neighborhoods U^*. With each point Q on B' let there be associated an arbitrarily small open neighborhood W^0 of Q on B admissibly represented by parameters (u). Let W denote the domain of points (u, s) for which (u) is on W^0 and $0 \leqq s < s_2$.

If Q is not a critical point of f^0 at least one of the partial derivatives $\partial H/\partial u^i$ does not vanish at Q. These derivatives are independent of s_1 and s_2. Accordingly the neighborhood W^0 of Q and the constant s_2 can be chosen so small that at least one of the derivatives $\partial H/\partial u^i$ does not vanish on the corresponding domain W.

If Q is a critical point of f^0 on B' it follows from the absence of critical points of f on B that at Q, $H_\nu = f^0_\nu \neq 0$. Moreover at Q, $H_\nu = f^0_\nu > 0$, for otherwise Q would be on $B - B'$. But H_ν depends upon s_1 only through the term Ls^2/s_1^2. Hence a diminution of s_1 merely increases the value of H_ν. Accordingly, the neighborhood W^0 of Q and the constant s_2 can be chosen so small that H_ν will not vanish on the corresponding domain W.

It now follows from an application of the Heine-Borel theorem that B' can be covered with a finite number of open neighborhoods W^0_h $(h = 1, \cdots, \lambda)$ such as W^0 and that accordingly one choice of the constant s_2 can be made such that on each of the corresponding domains $W = W_h$ at least one of the partial derivatives of H does not vanish. The function $H(x)$ has no critical points on these domains W_h and the lemma holds for the above choice of s_2.

17. **The type numbers of the critical sets of** H **neighboring** B. We shall prove the following lemma.

LEMMA 17.1. *The critical points of the function* $H(x)$ *other than those of* $f(x)$ *may be grouped into critical sets* σ^* *which correspond in a one-to-one manner to the complete critical sets* σ^0 *of* f^0 *on the negative boundary of* Σ. *The first n type numbers of the sets* σ^* *are equal to the corresponding type numbers of the corresponding sets* σ^0. *The* $(n + 1)$*-th type number of each critical set* σ^* *is zero.*

The first statement in the lemma is merely a rephrasing of the results obtained in §16. We proceed to the proof of the second statement.

Let σ^0 be a complete critical set of f^0 on the negative boundary of Σ. Let σ^* be the critical set of $H(x)$ which corresponds to the set σ^0 in the sense of §16. Recall that s^* is the value of s at points of σ^*. Let B^* be the manifold defined by $s = s^*$. We define a deformation J which deforms the points neighboring B^* through such points onto B^*. If $s(p)$ is the value of s at a point p neighboring B^*, the point p shall be deformed under J along the normal to B through p in such a fashion that as the time increases from 0 to 1 the difference $| s(p) - s^* |$ decreases to zero at a rate equal to its initial value.

Let c^0 be the value assumed by f^0 on σ^0 and let c^* be the value assumed by $H(x)$ on σ^*. We shall show that J deforms points neighboring σ^* below c^* through such points. Recall that

$$H(x) = f^0 + \left(G^0_r s + \frac{Ls^3}{3s_1^2}\right) \qquad (0 \leq s < s_2). \tag{17.1}$$

Among points neighboring σ^* the parenthesis in (17.1) depends only upon s and assumes a relative minimum value, say c^1, on σ^*. Let P be a point neighboring σ^* and Q its projection on B. Let f^0_Q be the value of f^0 at Q. Suppose that $H(x) < c^*$ at P. Observe that $c^* = c^0 + c^1$, and hence $f^0_Q < c^0$. During the deformation J, Q remains fixed, while the function $H(x)$ decreases monotonically to the value $H = f^0_Q + c^1$. Hence J has the property stated.

Let f^* be the function defined by H on B^*. Then $f^* = f^0 + c^1$. Hence f^* is a function of the same type as f^0. The set of points σ^* is a critical set of f^*. Let us denote σ^*, thought of as a critical set of f^*, by ω^*. Since f^* and f^0 differ by only a constant, it is clear that the type numbers of ω^* are equal respectively to the first n type numbers of σ^0.

Let VW be an admissible pair of neighborhoods of σ^* on R. For V sufficiently small and corresponding to V for W sufficiently small, the final images V' and W' of V and W respectively under the deformation J form an admissible pair of neighborhoods of ω^* on B^*. Since J deforms points below c^* through such points, it follows that maximal sets of spannable and critical k-cycles adm VW $(k = 0, 1, \cdots, n - 1)$ on R are deformed into maximal sets of spannable and critical k-cycles adm $V'W'$ on B^*.

Thus the first n type numbers of σ^* are equal respectively to those of ω^* and hence to those of σ^0. The type number m_n of σ^* is null, since H does not assume a maximum on σ^*. [8, p. 166.]

The proof of Lemma 17.1 is complete.

18. **Proof of Theorem** 13.1. The function $H(x)$ of the preceding sections has a positive normal directional derivative on the boundary B_1 of Σ_1. The spread B_1 is of class C^2 while the spread B is of class C^3. Were B_1 of class C^3, we could apply Theorem 10.2 to establish Theorem 13.1. We avoid the difficulty by transforming Σ_1 into Σ and B_1 into B in accordance with the following lemma.

Lemma 18.1. *There exists a one-to-one non-singular transformation T^* of class C^2 which carries Σ_1 into Σ and Σ_2 into a subdomain of Σ_2.*

The transformation T^* may be defined in essentially the same way as the transformation T^* in [9, §23] was defined.

We come to the proof of the Fundamental Lemma of §13. Suppose that under the transformation T^* the region Σ_2 is carried into a region Σ' and that the point (y) on Σ_2 is carried into the point (x) on Σ'. On Σ' we define a function $F(x)$ by the identity

$$F(x) \equiv H(y).$$

It follows from the properties of $H(x)$ enumerated in Lemmas 15.1 and 17.1 and the properties of T^* that the Fundamental Lemma is valid for the function $F(x)$.

Theorem 13.1 then follows as stated.

REFERENCES

[1] G. D. Birkhoff, and M. Hestenes, *The generalized minimax principle in the calculus of variations*, Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 413–432.

[2] A. B. Brown, *Critical sets of an arbitrary real analytic function of n variables*, Annals of Mathematics, vol. 32 (1931), pp. 512–520. Other papers by Brown are listed.

The definition of type numbers given in [2] is a modification of an earlier definition independently discovered by Brown. This definition of Brown made the type numbers depend upon more than the neighborhoods of the critical sets and was modified at the suggestion of Professors Morse and Lefschetz, as stated by Brown in a note in the above paper. These papers of Brown included important confirmations and extensions of the earlier work of Morse.

[3] S. S. Cairns, *On the cellular subdivision of n-dimensional regions*, Annals of Mathematics, vol. 33 (1932), pp. 671–680.

[4] S. Lefschetz, *Topology*, American Mathematical Society Colloquium Publications, New York, 1930.

[5] S. Lefschetz, *On critical sets*, this journal, vol. 1 (1935), pp. 392–412.

[6] M. Morse, *The analysis and analysis-situs of regular n-spreads in $(n + r)$-space*, Proceedings of the National Academy of Sciences, vol. 13 (1927), pp. 813–817.

[7] M. Morse, *The critical points of a function of n variables*, Transactions of the American Mathematical Society, vol. 33 (1931), pp. 72–91.

[8] M. Morse, *The Calculus of Variations in the Large*, American Mathematical Society Colloquium Publications, New York, 1934, Chapter VI.

[9] M. Morse, and G. B. Van Schaack, *The critical point theory under general boundary conditions*, Annals of Mathematics, vol. 35 (1934), pp. 545–571.

[10] M. Morse, and G. B. Van Schaack, *Abstract critical sets*, Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 258–263.

[11] M. Morse, *The critical points of a function of n variables*, Proceedings of the National Academy of Sciences, vol. 16 (1930), pp. 777–779. This paper contains the first definition of the type numbers of a critical set in the general case.

[12] M. Morse, *Functional topology and abstract variational theory*, Proceedings of the National Academy of Sciences, vol. 22 (1936), pp. 313–319.

THE INSTITUTE FOR ADVANCED STUDY.

PROOF THAT EVERY POSITIVE INTEGER IS A SUM OF FOUR INTEGRAL SQUARES

By R. D. Carmichael

The proof here given of the named classical theorem is a little longer than that offered by L. E. Dickson[1] in 1924, but has some elements of interest on account of the elegance of the method. Moreover, the reciprocal relations employed are of interest in themselves. Of the known proofs ours is most closely related to those of Dickson and Euler.

Let us write[2]

$$a^2 + \alpha b^2 + \beta c^2 + \alpha\beta d^2 = pq, \tag{1}$$

where $a, b, c, d, \alpha, \beta, p, q$ are integers, p, q are positive and α, β are not negative. Let $x, y, z, t, \lambda, \mu, \rho, \sigma$ be integers, and write

$$\begin{aligned} Aq = \{q\lambda + (ax - \alpha by - \beta cz - \alpha\beta dt)\}^2 + \alpha\{q\mu + (bx + ay + \beta dz - \beta ct)\}^2 \\ + \beta\{q\rho + (cx - \alpha dy + az + \alpha bt)\}^2 + \alpha\beta\{q\sigma + (dx + cy - bz + at)\}^2. \end{aligned} \tag{2}$$

The sum of the squares of the parenthesis quantities within the braces, multiplied by the indicated outside factors, is

$$(a^2 + \alpha b^2 + \beta c^2 + \alpha\beta d^2)(x^2 + \alpha y^2 + \beta z^2 + \alpha\beta t^2), \tag{3}$$

or $pq(x^2 + \alpha y^2 + \beta z^2 + \alpha\beta t^2)$, in accordance with the usual (and readily verified) product theorem for the forms in question. Hence A is an integer, and we have

$$\begin{aligned} A = q(\lambda^2 + \alpha\mu^2 + \beta\rho^2 + \alpha\beta\sigma^2) + p(x^2 + \alpha y^2 + \beta z^2 + \alpha\beta t^2) \\ + 2a(\lambda x + \alpha\mu y + \beta\rho z + \alpha\beta\sigma t) + 2\alpha b(-\lambda y + \mu x + \beta\rho t - \beta\sigma z) \\ + 2\beta c(-\lambda z - \alpha\mu t + \rho x + \alpha\sigma y) + 2\alpha\beta d(-\lambda t + \mu z - \rho y + \sigma x). \end{aligned} \tag{4}$$

The expression for A is invariant under the transformation

$$(p, q)(x, \lambda)(y, \mu)(z, \rho)(t, \sigma)(a, a)(b, -b)(c, -c)(d, -d)(\alpha, \alpha)(\beta, \beta). \tag{5}$$

So is equation (1). Hence we may perform on (2) the transformation (5) to obtain the relation

$$\begin{aligned} Ap = (px + a\lambda + \alpha b\mu + \beta c\rho + \alpha\beta d\sigma)^2 + \alpha(py - b\lambda + a\mu - \beta d\rho + \beta c\sigma)^2 \\ + \beta(pz - c\lambda + \alpha d\mu + a\rho - \alpha b\sigma)^2 + \alpha\beta(pt - d\lambda - c\mu + b\rho + a\sigma)^2. \end{aligned} \tag{6}$$

It may also be verified directly that (6) is implied by (1) and (4). In the presence of (1) equations (2) and (6) are equivalent. They constitute the *reciprocal relations* on which our proof is based.

Received January 9, 1936.

[1] L. E. Dickson, Amer. Journ. of Math., vol. 46 (1924), pp. 1–16; see esp. pp. 2–5.

[2] We here need the immediately following results (two paragraphs) only for $\alpha = \beta = 1$, but it seems well to put the more general formulas on record.

We shall now prove the following theorem.

I. *If p is a prime factor of the sum of four integral squares and is not a factor of each of them, then p is itself a sum of four integral squares.*

Since $2 = 1^2 + 1^2 + 0^2 + 0^2$, we assume in the following proof that p is an odd prime.

By hypothesis, there is some multiple pm of p which is a sum of four integral squares not all divisible by p. Then let pq be the *least* positive integral multiple of p which is a sum of four integral squares not all divisible by p. We then have a relation of the form

$$a^2 + b^2 + c^2 + d^2 = pq, \tag{7}$$

where a, b, c, d are integers not all divisible by p. We are to establish the theorem by proving that $q = 1$.

That $q < p$ follows at once from the fact that if a, b, c, d are replaced by their residues, modulo p, of least absolute value we have an equation of the character of (7) with $q < p$.

The numbers a, b, c, d in (7) have the greatest common divisor 1, since by dividing through by the square of their greatest common divisor we could otherwise replace (7) by another equation of like character and with a smaller value of q, contrary to the hypothesis that q has the least value possible.

We next show that we are led to a contradiction if we assume (as we now do for the moment) that $q > 2$. Since a, b, c, d have the greatest common divisor 1, integers x, y, z, t exist such that

$$ax - by - cz - dt = 1.$$

We employ these integers x, y, z, t. Use equation (2) with $\alpha = \beta = 1$, taking $\lambda = 0$ and choosing μ, ρ, σ so that the bases of the squares, after the first, in the second member of (2) are in absolute values not greater than $\frac{1}{2}q$. Since $q > 2$, we have $1 < \frac{1}{2}q$. Therefore, A in equation (2) is positive and is less than q if $q > 2$. Then from (6) we have an equation of the form

$$a_1^2 + b_1^2 + c_1^2 + d_1^2 = pA \qquad (0 < A < q < p)\text{.}$$

Since $A < p$, it follows that not all the numbers a_1, b_1, c_1, d_1 are divisible by p. Since $A < q$, we have a contradiction on assuming that $q > 2$. Hence $q \leqq 2$.

But if we assume that $q = 2$ we have

$$a^2 + b^2 + c^2 + d^2 = 2p.$$

Since p is odd, the numbers a, b, c, d are not all even. Then at least two of them, say a and b, are odd, and the other two are both odd or both even. In either case we may write

$$p = \{\tfrac{1}{2}(a + b)\}^2 + \{\tfrac{1}{2}(a - b)\}^2 + \{\tfrac{1}{2}(c + d)\}^2 + \{\tfrac{1}{2}(c - d)\}^2,$$

thus representing p as a sum of four integral squares. It is obvious that they are all prime to p. Hence $q \neq 2$.

It follows therefore that $q = 1$, and we have

$$p = a^2 + b^2 + c^2 + d^2,$$

a sum of four integral squares, as was to be proved.

We show next that every prime p satisfies the hypothesis in Theorem I, by giving the usual proof of the following long-known theorem:

II. *Every prime p is a divisor of a number of the form $u^2 + v^2 + 1$, where u and v are integers.*

For $p = 2$ take $u = 1, v = 0$. Henceforth, in the proof, assume that p is odd, say $p = 2k + 1$. Give to u successively the values $0, 1, 2, \cdots, k$. The least non-negative remainders of the corresponding numbers u^2, after division by p, are $k + 1$ in number and are all different since if two of them, say those corresponding to u_1 and u_2, are equal, we have $u_1^2 - u_2^2$, and hence $u_1 - u_2$ or $u_1 + u_2$ divisible by the prime p, $p = 2k + 1$, and this is impossible when u_1 and u_2 take values from the set $0, 1, 2, \cdots, k$. Likewise from $-v^2 - 1$ we obtain $k + 1$ different least non-negative remainders by giving to v the values $0, 1, 2, \cdots, k$ and dividing by p. Some number in one of these two sets, each of $k + 1$ remainders, must be equal to a number in the other set, since otherwise we would have $2k + 2$ different remainders all in the set $0, 1, 2, \cdots, 2k$, and this is impossible. For the corresponding numbers u and v we have $u^2 + v^2 + 1$ divisible by p. Hence the theorem is established.

Now the number 1 is a sum of four integral squares: $1 = 1^2 + 0^2 + 0^2 + 0^2$. From Theorems I and II it follows that every prime is a sum of four integral squares. Hence if there is any positive integer which is not the sum of four integral squares, it must be composite. If we suppose m to be the least positive integer which is not a sum of four integral squares, then we may write $m = m_1m_2$, $0 < m_1 < m$, $0 < m_2 < m$, whence it follows that m_1, and likewise m_2, is a sum of four integral squares. From the product theorem for a sum of four squares by a sum of four squares (implied in our derivation of the reciprocal relations) we see that m itself must be a sum of four integral squares, contrary to the hypothesis that m is the least positive integer which is not a sum of four integral squares. Thus we have the following theorem.

III. *Every positive integer is a sum of four integral squares.*

UNIVERSITY OF ILLINOIS.

TOPOLOGICAL FOUNDATIONS IN THE THEORY OF CONTINUOUS TRANSFORMATION GROUPS

By P. A. Smith

1. **Introduction.** This paper contains an account of some elementary topology connected with the notion of continuous group of transformations. From the modern point of view the group structures which are studied in the classical Lie theory are frequently not groups at all and have group-like properties only in restricted regions[1] of definition. This circumstance is due partly to the nature of the analysis involved and partly to the topological inadequacy of the ordinary types of coördinate systems and is therefore unavoidable. As a result, the group concepts of the classical theory are necessarily rather nebulous. We have attempted, however, to bring these concepts into sharper being, to crystallize their topological properties by means of definitions formulated postulationally in the spirit of the modern theory. We have not tried to obtain the highest degree of generality, or abstraction; our object is rather to define as simply as possible the types of group structures that one actually encounters and to study some of their simplest topological properties. We shall appeal occasionally to results in the Lie theory, but for the most part we have made it a point to treat situations which are essentially topological with topological methods.[2]

After the preliminary definitions we consider the question whether partial structures are essentially more general than total or completely defined structures or whether on the contrary every partial structure can be considered as being simply a piece of some total structure. In the most general case the answer is as yet unknown. We have, however, settled the question in certain special cases. In this connection we require a preliminary examination of the notion of transitivity, particularly transitivity in the neighborhood of a point, and this in turn requires the study of the topology of spaces whose elements are cosets of a given partial subgroup. In the final section we obtain new relations between the fundamental group of a given continuous group $\mathfrak{G}$ and that of a space $\mathfrak{X}$ in which $\mathfrak{G}$ operates transitively. Some results of this sort have already been obtained by Cartan;[3] our relations, however, have a more quantitative character, since they involve the ranks of certain subgroups of the fundamental

Received April 26, 1935.

[1] For this reason expositions of the Lie theory which do not define constantly the regions in which the objects dealt with have their existence would appear to have little validity beyond a purely formal analytic one. This systematic preoccupation with domains of definition seems only to occur in the classic treatise of Lie and Engel [9].

[2] In the study of Lie groups the advantages which are gained from an interplay between topology and analysis are excellently illustrated in Cartan's monograph [2]. We shall draw frequently from the ideas suggested in this work.

[3] [2], p. 27. See also Ehresmann [3], p. 399.

groups and the dimensions of $\mathfrak{G}$ and $\mathfrak{X}$. They are essentially an application of the results of a recent paper of the author [13].

2. **Notation and use of terms.** By a *space* we shall in general mean a Hausdorff space which satisfies the first denumerability axiom. This axiom asserts that there exists a complete system of neighborhoods which is of the form $\{V_n(p)\}$ $(n = 1, 2, \cdots)$, p being an arbitrary point of the space.

Let S be a space and q a point in S. We shall say that the sequence A_1, $A_2, \cdots$ of open sets in S *closes down* on q if the following conditions are satisfied: (1) $A_1 \supseteqq A_2 \supseteqq \cdots$ and (2) every point sequence $a_1, a_2, \cdots$ such that $a_i \in A_i$ converges to q. For example, let $\{V_i(p)\}$ be a complete system of neighborhoods for S and let $A_n = \prod_{i=1}^{n} V_i(q)$; then obviously the sequence $A_1, A_2, \cdots$ closes down on q.

We shall denote the intersection of the sets H and K by $H \wedge K$ and the closure of H by $\bar{H}$. The letters $a, b, \cdots, l, m, a_1, a_2, \cdots$ will consistently denote points in a space $\mathfrak{G}$ and $x, y, z, u, \cdots$ will denote points in a space $\mathfrak{X}$. $A, B, \cdots$ are open sets in $\mathfrak{G}$ and $X, Y, \cdots$ are open sets in $\mathfrak{X}$.

Group structures

3. Among the points $a, b, \cdots$ of a space $\mathfrak{G}$ we shall introduce "products" of the form ab and identify some or all of them with points of $\mathfrak{G}$ according to various assumptions to be considered. If ab is a point of $\mathfrak{G}$, we shall say that "the product ab is defined". For convenience, we shall refer to $\mathfrak{G}$ together with its product definitions as a *structure*.

Suppose that in the structure $\mathfrak{G}$ there exist open sets $A_1, \cdots, A_n$ $(n \geqq 2)$ such that for $h = 2, 3, \cdots, n$ the products

$$a_{h-1}a_h, \quad a_{h-2}(a_{h-1}a_h), \cdots, a_1(a_2(\cdots(a_{h-1}a_h)\cdots)),$$
$$a_ha_{h+1}, \quad (a_ha_{h+1})a_{h+2}, \cdots, (\cdots(a_ha_{h+1})a_{h+2}\cdots)a_n$$

(where a_i is an arbitrary point of A_i) are defined. The sets A_i then determine a *system* $(A_1, \cdots, A_n)$ *of degree* n for $\mathfrak{G}$. To assert that (A, B) is a system merely implies that every ab $(a \in A,\ b \in B)$ is identified with a point of $\mathfrak{G}$. If (A, B, C) is a system, then ab, bc, $a(bc)$, $(ab)c$ are points of $\mathfrak{G}$, and so on. If $\mathfrak{G}$ possesses a system of degree n, we shall say that $\mathfrak{G}$ is of degree $\geqq n$. If $\mathfrak{G}$ is of degree $\geqq k$ for each $k \geqq 2$, then we say that $\mathfrak{G}$ is of *infinite degree*. If $(A_1, \cdots, A_n)$ is a system in $\mathfrak{G}$, it is clear that $(A_h, A_{h+1}, \cdots, A_k)$ $(1 \leqq h < k \leqq n)$ is also a system in $\mathfrak{G}$.

4. We introduce the assumption

I. If ab, bc, $a(bc)$, and $(ab)c$ are defined, then $a(bc) = (ab)c$.

Suppose the structure $\mathfrak{G}$ satisfies I and is of degree $\geqq 3$; then $\mathfrak{G}$ possesses a system (A, B, C), and hence (§3) systems (A, B) and (B, C). The totality of

points ab ($a \epsilon A, b \epsilon B$) will be denoted by AB, and BC is similarly defined. The products $a(bc)$ ($c \epsilon C$) and $(ab)c$ are defined because of the definition of system, and they are identical by I. The point $a(bc) = (ab)c$ will be denoted by abc and the totality of these points by ABC.

In case $\mathfrak{G}$ is of degree $\geqq 4$, there is a system (A, B, C, D) and therefore (A, B), (B, C), (C, D), (A, B, C), (B, C, D) are systems and therefore the sets AB, BC, etc., are defined. In addition, it follows immediately from I and the definition of system that

$$(1) \qquad a(bcd), \qquad (abc)d, \qquad (ab)(cd), \qquad a(bc)d$$

are defined and are identical. We denote the point (1) by $abcd$ and the totality of these points by $ABCD$. In what follows we shall have to deal largely with systems of degrees 2, 3 and 4, contained in various types of structures satisfying I.

5. We introduce the further assumptions:

II. ab and a_1b can be identical points of $\mathfrak{G}$ only if $a = a_1$. ab and ab_1 can be identical only if $b = b_1$.

III. If $a_n \to a$ and $b_n \to b$ in $\mathfrak{G}$, and if $ab, a_1b_1, a_2b_2, \cdots$ are defined, then $a_nb_n \to ab$.

A structure $\mathfrak{G}$ which is of degree $\geqq 2$ and which satisfies I, II, III will be called a (continuous)[4] *semi-group*.[5]

A semi-group can be of infinite degree. For example, if there exists a system of the form $(\mathfrak{G}, \mathfrak{G})$, then obviously there are systems of the form $(\mathfrak{G}, \cdots, \mathfrak{G})$ and of arbitrarily high degree. A simple example of a structure which is a semi-group of this type, but which is not a group, is the following: let $\mathfrak{G}$ be the real numbers in the open interval (0, 1) and let ab be the arithmetic product of a and b. The characteristic property here is that all products are defined. In general, however, not all products need be defined, even though $\mathfrak{G}$ be of infinite degree and every point of $\mathfrak{G}$ be contained in some system.

6. Consider now the additional assumptions:

IV. $\mathfrak{G}$ contains an identity, that is, a point e such that $ae = ea = a$ for every a in $\mathfrak{G}$.

V. $\mathfrak{G}$ satisfies IV and possesses a system (E, E) where $e \epsilon E$.

VI. $\mathfrak{G}$ satisfies IV and there is an open set H containing e such that each a in H has a (right and left) inverse a^{-1} in $\mathfrak{G}$. If $a_n \to a$ in H, then $a_n^{-1} \to a^{-1}$.

A structure $\mathfrak{G}$ which satisfies assumptions I, III, IV, V, VI and for which $E = H = \mathfrak{G}$ is a (continuous) group; if $\mathfrak{G}$ satisfies these assumptions but is not a group, it will be called a *partial group*. In both cases the identity is unique, and inverses are unique when they exist. A semi-group, as we have defined it, may be a group or partial group; conversely, every group is a semi-group, since it

[4] Since we shall consider only structures which are continuous, that is, which satisfy III or similar assumptions, the word *continuous* will generally be omitted in the text.

[5] Cf. Eisenhart [4], p. 15.

obviously satisfies II. A partial group, however, need not be a semi-group although (§9) there is always a neighborhood of e in which II holds. A structure which is a group, partial group or semi-group will be called a *group structure*.

7. Let $\mathfrak{G}$ be a group or partial group and let $G_1, G_2, \cdots$ be a sequence of neighborhoods closing down (§2) on e and all contained in $E \wedge H$. Since $G_n \subseteq E$, the product G_nG_n is completely defined as a point set in G for every n. Since $G_n \subseteq H$, each a in G_n has an inverse a^{-1} and the totality of inverses will be denoted by G_n^{-1}.

For every neighborhood A of e, there is an N such that

$$G_nG_n \subseteq A, \qquad G_n^{-1} \subseteq A \qquad (n > N).$$

If this were not the case, for $n = 1, 2, \cdots$ there would be points a_n, b_n in G_n such that either $a_nb_n \bar{\epsilon} A$ or $a_n^{-1} \bar{\epsilon} A$ or both. But this is impossible because $a_nb_n \to e$ and $a_n^{-1} \to e$, since $a_n \to e$, $b_n \to e$.

Groups and partial groups are structures of infinite degree. Let n be chosen so large that $G_nG_n \subseteq E$. Then if a, b, c are in G_n, ab and bc are points in E and hence $a(bc)$ and $(ab)c$ are defined. Hence (G_n, G_n, G_n) is a system of degree 3. Now choose m so large that $G_m G_m \subseteq G_n$. If a, b, c, d are points of G_m, since $G_m \subseteq G_n$, ab, bc, etc.; $a(bc)$, $b(cd)$, etc., are defined and in fact are all contained in E. Hence $a(b(cd))$, etc., are defined and (G_m, G_m, G_m, G_m) is a system of degree 4. It is clear that this process yields systems of arbitrarily high degree.

8. There exists an N such that G_n^{-1} is an open set when $n > N$. We need only choose N so that $G_n^{-1} \subseteq H$ for $n > N$. For suppose that a^{-1} is not an inner point of G_n^{-1}. There exists a sequence $b_i \to a^{-1}$ such that $b_i \epsilon H$ and $b_i \bar{\epsilon} G_n^{-1}$. Then $b_i^{-1} \to a$ so that almost all the b_i's are in G_n. This is a contradiction. In similar manner, G_nG_n will be open if n is sufficiently large, likewise $G_nG_nG_n$, etc.

If we observe that the set $J_n = G_n + G_n^{-1}$ has the property that $J_n^{-1} = J_n$, we see readily from the preceding remarks that the sequence $G_1, G_2, \cdots$ can be chosen in such a way that $G_n^{-1} = G_n$ for each n.

9. *If $\mathfrak{G}$ is a group or partial group, there exists a neighborhood G of e such that* II *holds for all points a, b, a_1, b_1, in G.*

Proof. Let n be chosen so large that $G_nG_n \subseteq E$ and $G_n^{-1} \subseteq E$. We may take $G = G_n$. For if a, b, a_1 are in G_n and $ab = a_1b = c$, say, then $c \epsilon E$. Since $b^{-1} \epsilon E$, $(ab)b^{-1}$ and $(a_1b)b^{-1}$ are defined, and hence by I, $a = a_1$. Similarly, if $ab = ab_1$, we have $b = b_1$.

10. *A structure $\mathfrak{G}$ which satisfies* I, III, IV, V *and in which e has a neighborhood K which is an n-cell, is a group or partial group.*

Proof. Let the points of K be referred to a euclidean coördinate system and let $K_r \subseteq K$ be a spherical neighborhood of e with radius r, and J_r its boundary.

Let ρ be a value of r so small that $K_\rho + J_\rho \subset E$. On account of the continuity assumption III, we can choose $\rho_1 < \rho$ so small that

$$(1) \qquad d(a, ka) < \rho/2 \qquad (d = \text{distance})$$

whenever $a \epsilon J_\rho$ and $k \epsilon K_{\rho_1}$. Now the totality of points

$$kJ_\rho = \{ka, a \epsilon J_\rho\}$$

is a single-valued continuous image of J_ρ and on account of (1), kJ_ρ fails to contain e when $k \epsilon K_{\rho_1}$. Now let k be fixed in K_{ρ_1} and different from e. If a point k' moves from e to k along a radius, the successive images $k'J_\rho$ constitute a deformation of J_ρ to kJ_ρ and since none contains e, the looping coefficient[6] μ of kJ_ρ with respect to e equals that of J_ρ and hence $\mu \neq 0$. Moreover, if r varies from ρ to 0, the successive images kJ_r constitute a deformation of kJ_ρ, and since the final image is merely the point $ke = k \neq e$, its looping coefficient is 0. Hence some intermediate image $kJ_{\bar{r}}$ ($\rho > \bar{r} > 0$) contains e. This means that $e = k\bar{k}$, where $\bar{k}$ is a point of $J_{\bar{r}} \subset K_\rho$. Hence each point in K_{ρ_1} has a right-inverse in K_ρ. By the same reasoning there exists a K_{ρ_2}, each point of which has a right-inverse in K_{ρ_1}. Thus, if $a \epsilon K_{\rho_2}$, there is a point b in K_{ρ_1} such that $ab = e$, and a point b^{-1} in K_ρ such that $bb^{-1} = e$. Hence $a = a(bb^{-1}) = (ab)b^{-1} = eb^{-1} = b^{-1}$, so that $ba = e$ and b is also a left-inverse of a. Hence every point a in K_{ρ_2} has a left- and right-inverse a^{-1} in K_{ρ_1}; a^{-1} is unique, for if $aa^{-1} = aa' = e$, then $a^{-1}(aa^{-1}) = a^{-1}(aa')$ and $a^{-1} = a'$. Suppose that $a_n \to a$ in K_{ρ_2}. Since $a_n^{-1} \epsilon K_{\rho_1}$, we can choose a converging subsequence, say $a_{n_i}^{-1} \to c$, where $c \epsilon K_{\rho_1} + J_{\rho_1}$, and since $\rho_1 < \rho$, c is contained in E so that ac is defined. Hence $e = a_{n_i} a_{n_i}^{-1} \to ac$, so that $e = ac$, and in the same way $e = ca$. Hence $c = a^{-1}$, and $a_n^{-1} \to a^{-1}$. Therefore, if we take K_{ρ_2} for H, $\mathfrak{G}$ satisfies VI and is a group or partial group.

We may remark that a sufficient condition that $\mathfrak{G}$ be a group, or at least contain a substructure which is a group, is that there exist a system[7] of the form $(\mathfrak{G}, \mathfrak{G})$.

11. **Realization structures.** Let $\mathfrak{G}$ and $\mathfrak{X}$ be spaces. We introduce products of the form $a \cdot x$ ($a \epsilon \mathfrak{G}$, $x \epsilon \mathfrak{X}$), and identify some or all of them with points of $\mathfrak{X}$. This new type of structure will be denoted by $(\mathfrak{G}, \mathfrak{X})$. A product $a \cdot x$ which is identified with a point of $\mathfrak{X}$ is "defined".

Suppose that $\mathfrak{G}$ itself is a structure and possesses[8] a system $(A_1, \cdots, A_n)$ ($n \geqq 1$) and that X is an open set in $\mathfrak{X}$ such that

$$a_n \cdot x, \ a_{n-1} \cdot (a_n \cdot x), \ \cdots, \ a_1 \cdot (a_2 \cdot (\cdots (a_n \cdot x) \cdots)), \qquad (a_i \epsilon A_i, x \epsilon X)$$

are defined. There is thus defined a *realization system* $(A_1, \cdots, A_n; X)$ of degree n in $(\mathfrak{G}, \mathfrak{X})$. In particular, (A, X) is a realization system (of degree 1)

[6] For the definition and theory of looping coefficients, see Brouwer, [1].
[7] A proof can be found in the author's note [12].
[8] In case $n = 1$, (A_1) means merely the set A_1.

if and only if $a \cdot x$ is defined for each $a \in A$, $x \in X$. We shall have occasion later to consider realization systems of the form $(A_1, \cdots, A_n; x_0)$ when x_0 is a point in $\mathfrak{X}$.

We now introduce the assumptions

I_r. If ab, $b \cdot x$, and $a \cdot (b \cdot x)$ are defined, then so is $ab \cdot x$ (i.e., $(ab) \cdot x$), and $a \cdot (b \cdot x) = ab \cdot x$.

II_r. $a \cdot x$ and $a \cdot y$ can be identical points of $\mathfrak{X}$ only if $x = y$.

Note that if $\mathfrak{G}$ *is a group or partial group and* $(A, A; X)$ *a system in* $(\mathfrak{G}, \mathfrak{X})$ *such that* $e \in A$, *then* $e \cdot x = x$, *for every* $x \in X$. For let $e \cdot x = y$. Then $e \cdot (e \cdot x) = e \cdot y$ and $e \cdot x = e \cdot y$. Hence $y = x$ by II_r.

III_r. If $a_n \to a$ and $x_n \to x$ and if $a \cdot x, a_1 \cdot x_1, \cdots$ are defined, then $a_n \cdot x_n \to a \cdot x$.

12. Suppose that $\mathfrak{G}$ is a semi-group of degree $\geqq n$, and that the structure $(\mathfrak{G}, \mathfrak{X})$ satisfies I_r, II_r, III_r and possesses a realization system $(A_1, \cdots, A_m; X)$ of degree m $(2 \leqq m \leqq n)$. Then $(\mathfrak{G}, \mathfrak{X})$ is a (continuous) *realization* of degree $\geqq m$ of $\mathfrak{G}$. If $X = \mathfrak{X}$ we shall call $(\mathfrak{G}, \mathfrak{X})$ a *total* realization. If $\mathfrak{G}$ is a group or partial group, every realization is to satisfy the further condition that there be a system of the form $(A, A; X)$, where $e \in A$. It is easy to see (cf. §7) that realizations of groups and partial groups are of infinite degree. If $a_i \in A_i$, then since $a_h \cdot (a_{h+1} \cdot (a_{h+2} \cdots (a_n \cdot x) \cdots))$ is defined, it follows from I_r that $a_h a_{h+1} \cdot (a_{h+2} \cdots (a_n \cdot x) \cdots), \cdots, a_h a_{h+1} \cdots a_n \cdot x$ are defined and equal. We shall denote the totality of these points by $A_h A_h \cdots A_n \cdot X$.

13. Let $\mathfrak{G}$ be a group structure and (A, B, C) a system in $\mathfrak{G}$. Special realizations of $\mathfrak{G}$ can be defined as follows. Let $\mathfrak{X} = \mathfrak{G}$, and let $X = C$. If x is an arbitrary point of X, say $x = c$ $(c \in C)$, let $b \cdot x = bc$, $a \cdot (b \cdot x) = ab \cdot x = abc$. We have thus created a realization $(\mathfrak{G}, \mathfrak{X})$ with a system $(A, B; X)$ of degree 2. In similar fashion, we can form a realization of degree $\geqq n - 1$ if $\mathfrak{G}$ possesses a system of degree n. Realizations of the sort we have just described are usually called "first parameter groups".[9] The second parameter groups are defined as follows. First form the conjugate group structure $\mathfrak{G}^*$ with multiplication denoted by $a{*}b$, and defined by the relation $a{*}b = ba$ whenever ba is defined. If $\mathfrak{G}$ has a system of degree n, $(n \geqq 3)$, so does $\mathfrak{G}^*$ and the corresponding first parameter group of $\mathfrak{G}^*$ is the "second parameter group" of $\mathfrak{G}$.

14. $(\mathfrak{G}^r, \mathfrak{X}^n)$ denotes a realization of a semi-group $\mathfrak{G}^r$, where the spaces $\mathfrak{G}^r$ and $\mathfrak{X}^n$ are euclidean spaces of r and n dimensions. If (A, B) is a system in $\mathfrak{G}^r$, the set AB is open. For let a_0, b_0 be points in A, B and let $V \subsetneqq B$ be a bounded neighborhood of b_0. The correspondence $b \to a_0 b$ is $(1, 1)$ and continuous

[9] The definitions of first parameter group in the literature make no reference to the degree of $\mathfrak{G}$ and therefore, strictly speaking, do not define; for the first parameter group does not exist unless $\mathfrak{G}$ is of degree $\geqq 3$. This criticism does not apply, of course, when $\mathfrak{G}$ is explicitly assumed to be a group or partial group (i.e., to have an identity and inverses) since $\mathfrak{G}$ is then of infinite degree (§7).

because of II, III, and therefore when extended over the closed compact set $\bar{V}$ it is a homeomorphism. Hence $a_0\bar{V}$ is homeomorphic to $\bar{V}$ and hence a_0V is homeomorphic to V. From the invariance of regionality (Brouwer), a_0V is an open set and since $a_0b_0 \in a_0V \subseteq a_0B \subseteq AB$, a_0b_0 is an inner point of AB. But since a_0b_0 is an arbitrary point of AB, AB is open. By similar reasoning, $A \cdot X$ is an open set if (A, X) is a system in $(\mathfrak{G}^r, \mathfrak{X}^n)$, and G^{-1} is open if $\mathfrak{G}^r$ is a group or partial group and $G \subseteq H$ where H is defined in VI, §6.

As a consequence of the preceding remarks, we note that if (A, B, C) is a system in a group structure $\mathfrak{G}^r$ and if $(A, B; X)$ is a system in $(\mathfrak{G}^r, \mathfrak{X}^n)$, then (AB, C), (A, BC) are systems in $\mathfrak{G}^r$ and $(A, B; X)$, $(A; B \cdot X)$ are systems in $(\mathfrak{G}^r, \mathfrak{X}^n)$.

15. Structures in the classical theory. Let $\mathfrak{G}^r$ be a structure and $(\mathfrak{G}^r, \mathfrak{X}^n)$ a realization. We shall say that the systems (B, A) and $(A; X)$ in $\mathfrak{G}^r$ and $(\mathfrak{G}^r, \mathfrak{X}^n)$ are of class $C^{(n)}$ if

$$(ba)_i = \varphi_i(a_1, \cdots, a_r, b_1, \cdots, b_r) \qquad (i = 1, \cdots, r),$$

$$(a \cdot x)_i = f_i(x_1, \cdots, x_n, a_1, \cdots, a_r) \qquad (i = 1, \cdots, n),$$

where the φ's are defined for $a \in A$, $b \in B$ and the f's for $x \in X$, $a \in A$, and both sets of functions possess continuous derivatives of the first n orders throughout their regions of definition. The Lie theory of continuous groups deals largely with systems of degree 2 and of class $C^{(1)}$, $C^{(2)}$ or $C^{(3)}$, and systems of degree 3 in which the various systems of degree 2 which can be formed from such systems (§14) are of class $C^{(1)}$, $C^{(2)}$, or $C^{(3)}$.

Consider for example the first part of the "first fundamental theorem" of Lie. The statement of this theorem seems at first sight to imply only the existence of a system of the form $(A; X)$. An examination of details, however, shows that both the statement and proof[10] require a system of the form $(B, A; X)$. The theorem asserts, in fact, that if $(B, A; X)$ is a system in a representation $(\mathfrak{G}^r, \mathfrak{X}^n)$ of a semi-group $\mathfrak{G}^r$, and if the systems (B, A), $(A; X)$ and $(B; A \cdot X)$ are of class $C^{(1)}$, and if, moreover,

$$(1) \qquad \left| \frac{\partial (ba)_i}{\partial b_j} \right| \neq 0 \qquad (a \in A,\ b \in B),$$

$$\left| \frac{\partial (b \cdot x')_i}{\partial x'_j} \right| \neq 0 \qquad (b \in B,\ x' = a \cdot x \in A \cdot X = X', \text{ say}),$$

where the expressions on the left are functional determinants, the functions $(a \cdot x)_i = x'_i$ satisfy equations of the form

$$(2) \qquad \frac{\partial x'_j}{\partial a_i} = \sum_{k=1}^{r} \psi_{ik}(a)\xi_{kj}(x'),$$

[10] The proof to which we refer is that given in [9], vol. 1, p. 28.

where ψ's are defined throughout A and the ξ's throughout X' and where the determinant of the ψ's is not identically zero.

Suppose in particular that (B, A, C) is a system in a semi-group $\mathfrak{G}^r$ such that (B, A), (A, C) and (B, AC) are of class $C^{(1)}$ (so that (BA, C) is also of class $C^{(1)}$), and write $C = X$, $c = x$, $ac = a \cdot x$, thus forming a first parameter group (§13). Now since the transformations $b \to ba$ and $x' \to b \cdot x'$ are (1, 1) and all partial derivatives are continuous, there exist open sets A_1, B_1, X_1 contained in A, B, X such that relations (1) hold for $a \in A_1$, $b \in B_1$, $x' \in A_1 \cdot X_1$. Hence the functions $(a \cdot x)_i$ satisfy equations of the form (2) within A_1, X_1. Since the transformations $a \to a \cdot x$ are (1, 1), the determinant of the derivatives in (2) cannot vanish identically in A_1 for any x in X_1, and hence from (2) there exists a point $a_0 \in A_1$ such that

$$(3) \qquad | \psi_{ik}(a_0) | \neq 0 .$$

We now write

$$a \cdot x = t_a(x) \qquad (a \in A_1,\ x \in X_1).$$

Then t_a is a (1, 1) continuous transformation which carries x to $a \cdot x$. Let $\tau_a = t_a t_{a_0}^{-1}$, and $y = a_0 \cdot x$. Then

$$\tau_a(y) = \tau_a(a_0 \cdot x) = t_a(x) = a \cdot x \qquad (y \in a_0 \cdot X_1).$$

Hence the functions $(\tau_a(y))_i$ satisfy the equations (2) when substituted for x'_j. The family τ_a contains the identity, since $\tau_{a_0}(y) = y$.

Let us now assume that the systems (B, A) and $(B; A \cdot X)$ are of class $C^{(2)}$. An examination of the proof[10] of the Lie theorem shows that the functions $\psi(a)$ and $\xi(x')$ have continuous derivatives. Because of (3) and the fact that τ_{a_0} is the identity, it follows from the second part[11] of the Lie theorem that the functions $\tau_a(y)$ "define a group". What this means is that there exists a partial group $\mathfrak{G}^*$ defined in the same space as $\mathfrak{G}$ and in which a_0 is the identity and for which there is a realization system of the form $(A_2, A_2; Y_2)$, where $a_0 \in A_2 \subseteq A_1$, $Y_2 = a_0 \cdot X_2 \subseteq a_0 \cdot X_1$, and where A_2^{-1} is defined and $a \cdot y = \tau_a(y)$.

16. For later application (§32) we need to define certain additional functions. The transformation $t_{a_0}(x)$ is (1, 1) and continuous over X. Hence if X_3 is a bounded open set such that $\bar{X}_3 \subset X$, then $t_{a_0}(x)$ will be bi-continuous over X_3 so that $t_{a_0}^{-1}$ is continuous over $a_0 \cdot X_3$. Now choose X_3 such that $X_3 \subseteq X_2$, and let x_0 be a fixed point in X_3. Since $a_0 \cdot x_0 \in a_0 \cdot X_3$ and $a_0 \cdot X_3$ is an open set, we can choose a neighborhood $A'_2 \subseteq A_2$ of a_0 such that $a \cdot x_0 \in a_0 \cdot X_3$ when $a \in A'_2$. Hence there is a uniquely determined x in X_3 such that $a \cdot x_0 = a_0 \cdot x$. Hence we may write $x = \lambda(a)$, where $\lambda(a)$ is defined and single-valued over A'_2. Since $a \cdot x_0 \in a_0 \cdot X_0$, $t_{a_0}^{-1}(x')$ is defined and continuous whenever $x' = a \cdot x_0$ $(a \in A'_2)$. Hence we may write $\lambda(a) = t_{a_0}^{-1}(a \cdot x_0)$, and the transformation $a \to \lambda(a)$ is (1, 1)

[11] [9], vol. 3, p. 563. The proof given here holds when the ξ's and ψ's have continuous derivatives.

and continuous. Finally, let $\mu(a) = a \cdot x_0$ and let $A_3 \subseteqq A_2'$ be a neighborhood of a_0 such that the transformations $a \to \lambda(a)$ and $a \to \mu(a)$ are bi-continuous over A_3 (see the beginning of this section).

17. **The definition of group structures.** It is generally taken for granted in expositions of the Lie theory that if a family of transformations possesses the group property, there is determined a group structure for which the given transformations constitute a realization. So far as we are aware, this theorem has been never stated with complete precision, and a proof has been given only in case all the defining functions are analytic.[12] It seems therefore worth while to formulate a statement and proof which will justify an almost universal presumption.

Let $(\mathfrak{G}, \mathfrak{X})$ *be a structure in which the spaces* $\mathfrak{G}$ *and* $\mathfrak{X}$ *are euclidean.* Suppose that $a \cdot x$ and $b \cdot x$ are defined for all points x in some open X. We then write $a \doteq b$.

We introduce the assumptions

I′. If $b \cdot X$, $a \cdot (b \cdot X)$ and $c \cdot X$ are defined and if $c \cdot x = a \cdot (b \cdot x)$ for every $x \in X$, the same relation holds for every x for which $b \cdot x$, $a \cdot (b \cdot x)$ and $c \cdot x$ are defined.

II′. Each point a of $\mathfrak{G}$ possesses a neighborhood G_a within which the relation $a \doteq b$ can hold only if $a = b$.

III′. There exist open sets A, B, C, C being compact, and an open set X such that $B \cdot X$, $A \cdot (B \cdot X)$, $C \cdot X$ are defined, and for each a and b in A and B there are one or more points c in C such that $a \cdot (b \cdot x) = c \cdot x$ for every $x \in X$.

THEOREM. *If* $(\mathfrak{G}, \mathfrak{X})$ *satisfies* II_r, III_r *and* I′, II′, III′, *products of the form ab can be defined so that* $\mathfrak{G}$ *is converted into a semi-group for which* $(\mathfrak{G}, \mathfrak{X})$ *is a realization with a system of the form* $(A^*, B^*; X)$ *where* $A^* \subseteqq A$, $B^* \subseteqq B$.

Proof. Let a_0, b_0 be fixed points in A, B and let $\{c\}_{a_0b_0}$ be the set of c's in C such that $a_0 \cdot (b_0 \cdot x) = c \cdot x_0$ $(x \in X)$. This set is discrete; for if c_0 were a limit point, there could be no G_{c_0} satisfying II′. Since C is compact, $\{c\}_{a_0b_0}$ is finite, say $\{c\}_{a_0b_0} = \{c_1, \cdots, c_h\}$.

Let $F_1, \cdots, F_h$ be non-overlapping neighborhoods of $c_1, \cdots, c_h$ so small that $\bar{F}_i \subseteqq G_{c_i} \wedge C$ $(i = 1, \cdots, h)$. There exist neighborhoods A_0, B_0 of a_0, b_0 contained in A, B such that $\{c\}_{ab} \subset \Sigma F_i$ when $a \in A_0$, $b \in B_0$. If this were not the case, we could choose sequences $a_n \to a_0$, $b_n \to b_0$ in A_0, B_0 such that at least one point c_n' of $\{c\}_{a_nb_n}$ would fail to lie in ΣF_i. Since $c_n' \in C$ and C is compact, there is a converging subsequence $c_{n_i}' \to c_0' \in \bar{C}$. Now $a_{n_i}(b_{n_i} \cdot x) \to a_0 \cdot (b_0 \cdot x)$ and since $a_{n_i}(b_{n_i} \cdot x) = c_{n_i}' \cdot x$, we have $a_0 \cdot (b_0 \cdot x) = c_0' \cdot x$ so that $c_0' \in \{c\}_{a_0b_0}$; hence, infinitely many c_n''s lie in ΣF_i. This is a contradiction.

Let $a \times b$ be an arbitrary point in the topological product $A_0 \times B_0$. To each $a \times b$ there corresponds the set $\{c\}_{ab}$ in ΣF_i. Let H_i be those points of $A_0 \times B_0$ whose corresponding sets have at least one point in F_i. Then $A_0 \times B_0 = \Sigma H_i$. Moreover, H_i is *closed in* $A_0 \times B_0$. For suppose $a' \times b'$ is a limit point of H_1.

[12] [9], vol. 1, p. 16.

Choose a sequence $a_n \times b_n \to a' \times b'$. This implies that $a_n \to a'$, $b_n \to b'$. There exist points c_n in F_1 such that $a_n \cdot (b_n \cdot x) = c_n \cdot x$. We may choose a converging subsequence $c_{n_i} \to c' \epsilon \bar{F}_1 \subsetneq C$, and then, reasoning as above, we have $c' \cdot x = a' \cdot (b' \cdot x)$ so that $c' \epsilon \{c\}_{a'b'}$ and $a' \times b' \epsilon H_1$, and H_1 is closed. It follows that not all the sets H_i can be nowhere dense in $A_0 \times B_0$; we may suppose that H_1 contains an open set $H_{11} \subsetneq A_0 \times B_0$. We can choose open sets $A_{00} \subsetneq A_0$ and $B_{00} \subsetneq B_0$ such that $A_{00} \times B_{00} \subsetneq H_{11}$, and we shall have, for every $a \epsilon A_{00}$, $b \epsilon B_{00}$, a point c in H_{11} such that $a \cdot (b \cdot x) = c \cdot x$ $(x \epsilon X)$; c is unique, for if $a \cdot (b \cdot x) = c' \cdot x$ for every $x \epsilon X$, then $c \doteq c'$; but since c and c' are in G_{c_1}, $c = c'$.

Let a_{00}, b_{00} be fixed points in A_{00}, B_{00} and let $A^* = A_{00} \wedge G_{a_{00}}$, $B^* = B_{00} \wedge G_{b_{00}}$. The unique c which corresponds to every pair of points a, b in A^*, B^* we denote by ab. Thus $\mathfrak{G}$ is converted into a structure with a system (A^*, B^*) and for $(\mathfrak{G}, \mathfrak{X})$ there is a realization system $(A^*, B^*; X)$. *We shall now show that* $\mathfrak{G}$ *is a semi-group.*

We first verify I. Suppose lm, mn, $l(mn)$ and $(lm)n$ are defined. This implies that $l \epsilon A^*$, $m \epsilon B^*$; $m \epsilon A^*$, $n \epsilon B^*$; $lm \epsilon A^*$, $mn \epsilon B^*$. (This incidentally implies that $A^* \wedge B^* \neq 0$; hence I is trivially satisfied if A^* and B^* do not meet.) Let x be an arbitrary point of X and let $y = n \cdot x$. Then $m \cdot y$ is defined, since

$$mn \cdot x = m \cdot (n \cdot x) = m \cdot y. \tag{1}$$

Furthermore, $lm \cdot y$ is defined, since

$$(lm)n \cdot x = lm \cdot (n \cdot x) = lm \cdot y. \tag{2}$$

Finally, $l \cdot (mn \cdot x)$ is defined and from (1),

$$l \cdot (mn \cdot x) = l \cdot (m \cdot (n \cdot x)) = l \cdot (m \cdot y), \tag{3}$$

so that $l \cdot (m \cdot y)$ is defined. Since $l(mn \cdot x) = l(mn) \cdot x$, we have from (3)

$$l(mn) \cdot x = l \cdot (m \cdot y). \tag{4}$$

Since $lm \cdot y$ and $l \cdot (m \cdot y)$ are defined, we have from I′ $lm \cdot y = l \cdot (m \cdot y)$. Hence by (2) and (4), $l(mn) \cdot x = (lm)n \cdot x$, for every $x \epsilon X$. Hence $l(mn) = (lm)n$, since these points are both in G_{c_1}. Hence I is proved.

Suppose that $ab = a_1 b$. This implies that

$$a \cdot (b \cdot x) = a_1 \cdot (b \cdot x) \qquad (a, a_1 \epsilon A^*, b \epsilon B^*) \tag{5}$$

for every $x \epsilon X$. By the argument of §14, $b \cdot X$ is an open set and therefore (5) implies that $a \doteq a_1$. Since $A^* \subsetneq G_{a_{00}}$, we have $a = a_1$. Again, if $ab = ab_1$, then $a \cdot (b \cdot x) = a \cdot (b_1 \cdot x)$. Hence from II$_r$ $b_1 \cdot x = b \cdot x$. Hence $b_1 \doteq b$ and since $B^* \subsetneq G_{b_{00}}$, $b_1 = b$. This verifies II.

Finally, we establish continuity (III). Suppose $a_n \to a$ in A^* and $b_n \to b$ in B^*. Then $a_n \cdot (b_n \cdot x) \to a \cdot (b \cdot x)$ from III$_r$. Since $a_n b_n \epsilon \bar{F}_1 \subsetneq G_{c_1} \wedge C$, there is a converging subsequence $a_{n_i} b_{n_i} \to c \epsilon \bar{F}_1$, and hence $c \cdot x = ab \cdot x$. Since c and ab

are both in G_{e_1}, we have $c = ab$. Thus c is independent of the particular converging subsequence chosen, and $a_n b_n \to ab$.

Thus $\mathfrak{G}$ is a semi-group. $(\mathfrak{G}, \mathfrak{X})$ is obviously a realization of $\mathfrak{G}$ and our theorem is established.

Remarks. In Lie's proof of the existence of the function ab, the assumption II′ is replaced by the assumption that the "parameters are essential" and at the same time the functions which define $a \cdot x$ are assumed to be analytic; but it is not easy to see from the proof how the assumption of essentiality can be made to replace II′ in any situation short of the analytic one. To be sure, in certain cases, essentiality implies II′; but we can prove this only if the theorem in question is assumed proved, (§38). It is of course obvious that II′ can be omitted in our theorem if it is assumed in III′ that for every a, b there is a *unique* c. This seems to be a tacit assumption in most treatments.[13]

Transitivity

18. Let $(\mathfrak{G}, \mathfrak{X})$ be a realization of a group or partial group $\mathfrak{G}$. A system $(G, \cdots, G; x_0)$ $(e \in G)$ will be called a *transitive* system if $G \subseteq H$, $G^{-1} = G$, and if $G \cdot x_0$ is an open set.[14]

THEOREM. *If $(G, G, G; x_0)$ is a transitive system and if there is a system $(G; X)$ such that $GG \cdot x_0 \subseteq X$, then $GG \cdot x_0$ is an open set.*

Proof. Let $gg_1 \cdot x_0$ $(g, g_1 \in G)$ be an arbitrary point in $GG \cdot x_0$, and let x_n be an arbitrary sequence of points such that $x_n \to gg_1 \cdot x_0$. Since $gg_1 \cdot x_0 \in GG \cdot x_0 \subseteq X$, almost every x_n is in X and hence almost every product $g^{-1} \cdot x$ is defined (since $g^{-1} \in G^{-1} = G$). Hence,

$$g^{-1} \cdot x_m \to g^{-1}gg_1 \cdot x_0 = g_1 \cdot x_0 \qquad (m > \bar{n}, \text{ say}).$$

Since $g_1 \cdot x_0 \in G \cdot x_0$ and $G \cdot x_0$ is open, almost every $g^{-1} \cdot x_m$ is in $G \cdot x_0$. Hence almost every x_m is in $gG \cdot x_0$ and this last set therefore contains $gg_1 \cdot x_0$ as an inner point. This proves the theorem.

19. Let $\mathfrak{G}$ be a group and $(\mathfrak{G}, \mathfrak{X})$ a total (§12) realization of $\mathfrak{G}$, and $(G, G, G; x_0)$ a transitive system. It follows from the preceding theorem that $GG \cdot x_0$ is open. As a matter of fact, since the degree of a system no longer plays any part in the argument, it is clear that if G is any open set containing e, and if $G \cdot x_0$ is open, then $GG \cdot x_0$, $GGG \cdot x_0$, $\cdots$ are all open. Let $G^{(n)} = GG \cdots G$ (n factors). Then $G^{(1)} \subseteq G^{(2)} \subseteq \cdots$, and $G^{(1)} \cdot x_0 \subseteq G^{(2)} \cdot x_0 \subseteq \cdots$. Suppose that $\mathfrak{G}$ is connected. From a theorem of Schreier [10], $\lim_{n \to \infty} G^{(n)} = \mathfrak{G}$. Hence $\lim_{n \to \infty} G^{(n)} \cdot x_0 = \mathfrak{G} \cdot x_0$. Suppose further that $\mathfrak{G}$ is compact and $\mathfrak{X}$ is connected. *We assert that* $\mathfrak{G} \cdot x_0 = \mathfrak{X}$.

[13] There are various ways of stating assumptions which will lead to the existence of a group or partial group. We shall not, however, discuss these questions further at present nor shall we discuss the possible relations between the various determinations of the function ab in the case where it is multiple-valued.

[14] The last condition is the essential one; the first two are only a convenience.

Since $\mathfrak{G}\cdot x_0$ is open, we need only show that $\mathfrak{G}\cdot x_0$ is closed. Let $x_n \to x$, where $x_n \in \mathfrak{G}\cdot x_0$. We may write $x_n = g_n\cdot x_0$ $(g_n \in \mathfrak{G})$. Since $\mathfrak{G}$ is compact, there is a converging subsequence $g_{n_i} \to g$; then $g_{n_i}\cdot x_0 \to g\cdot x_0$, and hence $g\cdot x_0 = x$ and $x \in \mathfrak{G}\cdot x_0$. Hence $\mathfrak{G}\cdot x_0$ is closed.

As a consequence of the relation $\mathfrak{G}\cdot x_0 = \mathfrak{X}$, we note that for any pair of points x, y in $\mathfrak{X}$ there is a point c in $\mathfrak{G}$ such that $c\cdot x = y$, that is, the realization $(\mathfrak{G}, \mathfrak{X})$ is *transitive.* Since $x = a\cdot x_0$ and $y = b\cdot x_0$ for properly chosen a, b, then $ba^{-1}\cdot x = b\cdot x_0 = y$, and we may take $c = ba^{-1}$. We have the theorem

If $(\mathfrak{G}, \mathfrak{X})$ *is a total realization of a compact connected group, if* $\mathfrak{X}$ *is connected, and if for some* x_0 *and open* G $(e \in G)$ $G\cdot x_0$ *is open, then* $(\mathfrak{G}, \mathfrak{X})$ *is transitive.*

This theorem states, in effect, that complete transitivity is a consequence of transitivity in the neighborhood of a point. We shall now turn to the converse problem, and show what type of local transitivity is implied by complete transitivity. Our results will be used later in the proofs of certain imbedment theorems. Although we are able to avoid the assumption that $\mathfrak{G}$ is compact, an assumption which would make the proofs fairly obvious, it will be expedient to assume from now on that $\mathfrak{G}$ is separable, metric, and semi-compact, and that $\mathfrak{X}$ is metric and complete. The particular properties which we shall require and which follow from these assumptions are

(a) every open set G in $\mathfrak{G}$ can be expressed as the sum of a denumerable set of closed, compact sets;

(b) if $X_1 + X_2 + \cdots$ is an open set in $\mathfrak{X}$, not every X_i can be nowhere dense in ΣX_i.[15]

20. Suppose that $\mathfrak{G}$ is a group or partial group, $(\mathfrak{G}, \mathfrak{X})$ a realization. Suppose, further, that $(\mathfrak{G}, \mathfrak{X})$ possesses a system $(G, G, G, G; x_0)$ such that $e \in G$, $G^{-1} = G$. *All the points in* GG *have inverses.* In fact, if $g, h \in G$, then $g^{-1}, h^{-1} \in G$, since $G^{-1} = G$. The product $(gh)(h^{-1}g^{-1})$ is defined and equals e. Hence $(gh)^{-1} = h^{-1}g^{-1}$ and $(GG)^{-1} = G^{-1}G^{-1} = GG$.

We shall construct a new space[16] whose elements are certain subsets of G.

Let $\mathfrak{g}$ consist of all points g in GG such that $g\cdot x_0 = x_0$. We assert that

(1) $\mathfrak{g}\mathfrak{g} \wedge GG = \mathfrak{g}$.

(2) If $a\mathfrak{g} \wedge G = b\mathfrak{g} \wedge G$ $(a, b, \in G)$, then $b^{-1}a \in \mathfrak{g}$.

(3) If $h_1\cdot x_0 = h_2\cdot x_0$ $(h_1, h_2 \in G)$, then $h_1 \in h_2\mathfrak{g} \wedge G$.

(4) $(hg)\mathfrak{g} \wedge G = h\mathfrak{g} \wedge G$, if $h \in G$, $g \in \mathfrak{g}$, $hg \in G$.

(5) If $h_1 \in h\mathfrak{g} \wedge G$ $(h \in G)$, then $h_1\mathfrak{g} \wedge G = h_2\mathfrak{g} \wedge G$.

Proof. (1) $\mathfrak{g}\mathfrak{g}$ is defined because $\mathfrak{g} \subseteq GG$ and G is part of a system of degree 4. Since $\mathfrak{g}\mathfrak{g}\cdot x_0 = x_0$, then certainly $(\mathfrak{g}\mathfrak{g} \wedge GG)\cdot x_0 = x_0$, and hence $\mathfrak{g}\mathfrak{g} \wedge GG \subseteq \mathfrak{g}$. Obviously, $\mathfrak{g} \subseteq \mathfrak{g}\mathfrak{g} \wedge GG$; hence (1) is proved.

[15] See Hausdorff, [5], pp. 136–141, particularly X, p. 141.

[16] Cf. Cartan, [2], p. 26.

(2) Since $a \in a\mathfrak{g} \wedge G$ (because $e \in G$), then $a \in b\mathfrak{g} \wedge G$; that is, $a = bg$ where $g \in \mathfrak{g}$. Hence $b^{-1}a = g$.

(3) Obvious.

(4) An arbitrary point in $(hg)\mathfrak{g} \wedge G$ is of the form $(hg)g_1$, where $g_1 \in \mathfrak{g}$, $(hg)g_1 \in G$. Now $(hg)g_1 \cdot x_0 = (hg) \cdot (g_1 \cdot x_0) = hg \cdot x_0 = h \cdot (g \cdot x_0) = h \cdot x_0$. Hence, by (3), $(hg)g_1 \in h\mathfrak{g} \wedge G$. Hence $(hg)\mathfrak{g} \wedge G \subseteq h\mathfrak{g} \wedge G$. Conversely, consider an arbitrary point in $h\mathfrak{g} \wedge G$; it is of the form hg_2, where $g_2 \in \mathfrak{g}$ and $h\mathfrak{g}_2 \in G$. We have $hg_2 = (he)g_2 = (h(gg^{-1}))g_2 = ((hg)g^{-1})g_2$, since $hg \in G$, and since $g^{-1} \in GG$ and $g_2 \in GG$, it follows that $g^{-1}g_2$ is defined, so that $hg_2 = (hg)(g^{-1}g_2)$. Now $hg_2 \in G$ and $(hg)^{-1} \in G^{-1} = G$. Hence $(hg)^{-1}(hg_2)$ is defined, and $(g^{-1}g_2) = (hg)^{-1}(hg_2) \in GG$. Since $g^{-1}g_2 \in \mathfrak{g}\mathfrak{g}$, we have $g^{-1}g_2 = \mathfrak{g}\mathfrak{g} \wedge GG = \mathfrak{g}$ by (3). Hence $hg_2 = (hg)(g^{-1}g_2) \in hg\mathfrak{g}$, and since $hg_2 \in G$, we have $hg_2 \in hg\mathfrak{g} \wedge G$, and (4) is proved.

(5) Since $h_1 \in h_2\mathfrak{g} \wedge G$, we may write $h_1 = h_2g$, $g \in \mathfrak{g}$. Hence $h_1\mathfrak{g} \wedge G = h_2g\mathfrak{g} \wedge G = h_2\mathfrak{g} \wedge G$ by (4).

From (5) it follows that if $h_1\mathfrak{g} \wedge G$ and $h_2\mathfrak{g} \wedge G$ have a point in common, they are identical. Moreover, if h is a point of G, then $h \in h\mathfrak{g} \wedge G$. Hence *the points of G fall into a family of mutually exclusive sets* which we shall denote by $\mathfrak{h}_\xi$, $\mathfrak{h}_\eta$, $\cdots$. Let us choose in each of these sets, once and for all, a single point, denoting the point in $\mathfrak{h}_\xi$ by h_ξ. By (5) we have $\mathfrak{h}_\xi = h_\xi\mathfrak{g} \wedge G$.

21. Let $\mathfrak{H}$ be the totality of sets $\mathfrak{h}_\xi$, $\mathfrak{h}_\eta$, $\cdots$. We shall introduce a topology into $\mathfrak{H}$ as follows. Let $G_1, G_2, \cdots$ be a sequence of neighborhoods contained in G and closing down on e. Let

$$V_n(\mathfrak{h}_\xi) = (h_\xi G_n \wedge G)\mathfrak{g} \wedge G.$$

$V_n(\mathfrak{h}_\xi)$ obviously consists of a set of $\mathfrak{h}_\eta$'s and we may therefore think of $V(\mathfrak{h}_\xi)$ as a subset of $\mathfrak{H}$. The totality of these subsets is to constitute a complete set of neighborhoods for $\mathfrak{H}$. *We shall show that the Hausdorff axioms are satisfied.*

First, it is clear that if $V_n(\mathfrak{h}_\xi)$ and $V_m(\mathfrak{h}_\xi)$ are two neighborhoods of $\mathfrak{h}_\xi$ and if $k > \max(m, n)$, we have $V_k(\mathfrak{h}_\xi) \subseteq V_n(\mathfrak{h}_\xi) \wedge V_m(\mathfrak{h}_\xi)$.

Next, suppose $\mathfrak{h}_\xi \neq \mathfrak{h}_\eta$. We show that for n sufficiently large, $V_n(\mathfrak{h}_\xi) \wedge V_n(\mathfrak{h}_\eta) = 0$. If this were not so, we would have for each n points a_n, b_n in G_n such that $h_\xi a_n\mathfrak{g} \wedge G = h_\eta b_n\mathfrak{g} \wedge G$, $h_\xi a_n \in G$, $h_\eta b_n \in G$. Then $h_\xi a_n \in h_\eta b_n \wedge G$, and therefore we may write

$$h_\xi a_n = h_\eta b_n g_n \qquad (g_n \in \mathfrak{g}). \tag{1}$$

Hence $(h_\eta b_n)^{-1}(h_\xi a_n) = (h_\eta b_n)^{-1}(h_\eta b_n g_n)$. Since h_ξ, a_n, h_η, b_n are in G and $h_\eta b_n g_n \in GGG$, and since G is part of a system of degree 4, we have $(h_\eta b_n)^{-1}(h_\eta b_n g_n) = (h_\eta b_n)^{-1}(h_\eta b_n)g_n = g_n$. It follows from (1) that $h_\xi a_n \cdot x_0 = h_\eta b_n g_n \cdot x_0 = (h_n b_n) \cdot (b_n \cdot x_0) = h_\eta b_n \cdot x_0$, and on letting $n \to \infty$ and observing that $a_n \to e$, $b_n \to e$, we see that $h_\xi \cdot x_0 = h_\eta \cdot x_0$. Hence $h_\xi \in h_\eta\mathfrak{g} \wedge G$ by (3), §20, and by (5), §20, $h_\xi\mathfrak{g} \wedge G = h_\eta\mathfrak{g} \wedge G$; that is, $\mathfrak{h}_\xi = \mathfrak{h}_\eta$. This is a contradiction.

Finally, let $\mathfrak{h}_\eta$ be an element in $V_n(\mathfrak{h}_\xi)$. We shall show that for k sufficiently

large, $V_k(\mathfrak{h}_\eta) \subseteq V_n(\mathfrak{h}_\xi)$. We have $\mathfrak{h}_\eta \subset (\mathfrak{h}_\xi G_n \wedge G)\mathfrak{g} \wedge G$, and therefore we may write

$$(2) \qquad \mathfrak{h}_\eta = (h_\xi h_n)\mathfrak{g} \wedge G \qquad (h \in G_n,\ h_\eta h_n \in G)\ .$$

Let m be chosen so large that $h_n G_m \subseteq G_n$ and $h_\xi h_n G_m \subseteq G$. Now $\mathfrak{h}_\eta = h_\eta \mathfrak{g} \wedge G$, and on comparing this with (2) we have, from (2), §20, $(h_\xi h_n)^{-1} h_\eta = g$, say, where $g \in \mathfrak{g}$. Since $(h_\xi h_n)^{-1} \in G^{-1} = G$, we may write $h_\eta = (h_\xi h_n) g$. Hence

$$V_k(\mathfrak{h}_\eta) = (h_\eta G_k \wedge G)\mathfrak{g} \wedge G = (((h_\xi h_n)g)G_k \wedge G)\mathfrak{g} \wedge G\ .$$

Since $h_\xi h_n \in G$, $\mathfrak{g} \subseteq GG$, $G_k \subseteq G$ and G defines a system of degree 4, we have $(h_\xi h_n)gG_k = h_\xi h_n(gG_k)$. Let k be chosen so large that $gG_k \subseteq G_m g$. Then $((h_\xi h_n)g)G_k \subseteq h_\xi h_n(G_m g) = ((h_\xi h_n)G_m)g$ and $V_k(h_\eta) \subseteq (((h_\xi h_n)G_m)g \wedge G)\mathfrak{g} \wedge G$, and since $h_\xi h_n G_m \subseteq G$ by choice of m, we can apply (4), §20, to the expression on the right; it results that $V_k(\mathfrak{h}_\eta) \subseteq ((h_\xi h_n)G_m \wedge G)\mathfrak{g} \wedge G$. Since $(h_\xi h_n)G_m = h_\xi(h_n G_m)$ and this last set is contained (by the choice of m) in $h_\xi G_n$, we have $V_k(\mathfrak{h}_\eta) \subseteq (h_\xi G_n \wedge G)\mathfrak{g} \wedge G = V_n(\mathfrak{h}_\xi)$.

We have shown that $\mathfrak{H}$ is a Hausdorff space; obviously it satisfies the first denumerability axiom.

22. The fact that $\mathfrak{g}$ satisfies the relation $\mathfrak{g}\mathfrak{g} \wedge GG = \mathfrak{g}$ shows that $\mathfrak{g}$ is a sort of (partial) subgroup of $\mathfrak{G}$, and it is possible to start with this property, together with the assumption that $\mathfrak{g}$ is closed in GG, and construct $\mathfrak{H}$ without reference to a space $\mathfrak{X}$. In case $\mathfrak{G}$ is a group and $G = \mathfrak{G}$, $\mathfrak{g}$ is a closed subgroup, and the elements of $\mathfrak{H}$ are the left cosets $\mathfrak{g}$. The verification of the Hausdorff axioms is now quite simple. For example, suppose $\mathfrak{h}_\xi \neq \mathfrak{h}_\eta$. We have $\mathfrak{h}_\xi = h_\xi \mathfrak{g}$, $\mathfrak{h}_\eta = h_\eta \mathfrak{g}$, $V_n(h_\xi) = h_\xi G_n \mathfrak{g}$, $V_n(h_\eta) = h_\eta G_n \mathfrak{g}$. Suppose that for each n, $V_n(\mathfrak{h}_\xi) \wedge V_n(\mathfrak{h}_\eta) = 0$. We could then choose a_n, b_n in G_n such that $h_\xi a_n \mathfrak{g} = h_\eta b_n \mathfrak{g}$. Hence $(b_n h_\eta)^{-1}(h_\xi a_n) = g$, where $g \in \mathfrak{g}$. Since $a_n \to e$, $b_n \to e$, we have $h_n \to h_\eta^{-1} g_\xi$, and since $\mathfrak{g}$ is closed, $h_\eta^{-1} h_\xi \in \mathfrak{g}$. Hence $h_\xi = h_\eta g$ $(g \in \mathfrak{g})$ and $h_\xi \mathfrak{g} = h_\eta g \mathfrak{g} = h_\eta \mathfrak{g}$, or $\mathfrak{h}_\xi = \mathfrak{h}_\eta$. This is a contradiction.

One further remark concerning the present case ($G = \mathfrak{G}$ = a group): if $a_n \to a$, and if $a\mathfrak{g} = \mathfrak{h}_\xi$, $a_n \mathfrak{g} = \mathfrak{h}_{\xi_n}$, then obviously $\mathfrak{h}_{\xi_n} \to \mathfrak{h}_\xi$ in $\mathfrak{H}$. Conversely, if $\mathfrak{h}_\xi = a\mathfrak{g}$ and if $\mathfrak{h}_{\xi_n} \to \mathfrak{h}_\xi$, there can be chosen a sequence a_n such that $a_n \mathfrak{g} = \mathfrak{h}_{\xi_n}$ and $a_n \to a$. This is perhaps not quite obvious, but the proof offers no real difficulties, and the details closely resemble those in the proof of a later theorem, (§25).

23. Let us now return to the system $(G, G, G, G; x_0)$ of §20, and *assume that it is transitive*. Let the open set $G \cdot x_0$ be denoted by X_0, and let $\mathfrak{h}_x$ $(x \in X_0)$ consist of all points h in G such that $h \cdot x_0 = x$. If h_1 is one of these points, we have

$$(1) \qquad \mathfrak{h}_x = h_1 \mathfrak{g} \wedge G\ .$$

For, let $h_1 g$ $(g \in \mathfrak{g})$ be an arbitrary point of $h_1 \mathfrak{g} \wedge G$. Then $(h_1 g) \cdot x_0 = h_1(g \cdot x_0) = h_1 \cdot x_0 = x$. Hence $h_1 g \in \mathfrak{h}_x$ and $h_1 \mathfrak{g} \wedge G \subseteq \mathfrak{h}_x$. Conversely, if h is an arbitrary

point in $\mathfrak{h}_x$, since $h \cdot x_0 = x = h_1 \cdot x_0$, we have $h \in h_1\mathfrak{g} \wedge G$ by (3), §20. Hence $\mathfrak{h}_x \subseteq h_1\mathfrak{g} \wedge G$, and (1) is proved.

It follows that each $\mathfrak{h}_x$ is a $\mathfrak{h}_\xi$; conversely, it is obvious that each $\mathfrak{h}_\xi$ is a $\mathfrak{h}_x$. Thus the elements of $\mathfrak{H}$ can be denoted by $\mathfrak{h}_x, \mathfrak{h}_y, \cdots$ and the corresponding points $h_\xi, h_\eta, \cdots$ by $h_x, h_y, \cdots$. If $x \neq y$, then $h_x \neq h_y$, for otherwise we would have $x = h_x \cdot x_0 = h_y \cdot x_0 = y$. We conclude: there is a (1, 1) correspondence $x \to h_x$ between the points of $\mathfrak{H}$ and those of X_0; for each x, $\mathfrak{h}_x = h_x\mathfrak{g} \wedge G$.

We shall now show that the correspondence is bi-continuous. Let us denote $\mathfrak{h}_x$ by $\mathfrak{h}(x)$. We shall first show that if $\mathfrak{h}(x_m) \to \mathfrak{h}(x)$, then $x_m \to x$. Let V be an arbitrary neighborhood of x. Choose $\bar{n}$ so large that $h_xG_{\bar{n}} \cdot x_0 \subseteq V$. Almost every $\mathfrak{h}(x_m)$ is in $V_{\bar{n}}(\mathfrak{h}_x) = (h_xG_{\bar{n}} \wedge G)\mathfrak{g} \wedge G$, and is therefore of the form $h_x\, a_m\mathfrak{g} \wedge G$ $(a_m \in G_{\bar{n}}, h_xa_m \in G)$. Hence, for almost every m, $h_xa_m \cdot x_0 = x_m$, and therefore, by (1), almost every x_m is in V. Hence $x_m \to x$.

Conversely, if $x_m \to x$, then $\mathfrak{h}(x_m) \to \mathfrak{h}(x)$. To prove this we first establish the following

Lemma. *Let $x, x_1, x_2, \cdots$ be points in X_0 such that for each n, $h_x \cdot x_n \in X_0$. If $\mathfrak{h}(x_m) \to \mathfrak{h}(x_0)$, then $\mathfrak{h}(h_x \cdot x_n) \to \mathfrak{h}(h_x \cdot x_0)$, and conversely.*

Proof. Let $\bar{m}$ be chosen so large that $h_xG_{\bar{m}} \subseteq G$. Let m be an arbitrary integer $> \bar{m}$. Then $h_xG_m \subseteq h_xG_{\bar{m}} \subseteq G$. We have also $V_m\mathfrak{h}(x_0) = G_m\mathfrak{h}(x_0) \wedge G$, since $G_m \subseteq G$, $V_m\mathfrak{h}(h_x \cdot x_0) = V_m(\mathfrak{h}_x) = (h_xG_m \wedge G)\ \mathfrak{h}(x_0) \wedge G = h_x\, G_m\mathfrak{h}(x) \wedge G$.

Now suppose that $\mathfrak{h}(x_n) \to \mathfrak{h}(x_0)$. Then for almost all n $\mathfrak{h}(x_n) \in V_m(\mathfrak{h}(x_0)$ so that we may write

$$\mathfrak{h}(x_n) = c_n\mathfrak{h}(x_0) \wedge G \qquad (c_n \in G_m)\ ,$$

(n sufficiently large). Since $m > \bar{m}$, we have $h_xc_n \in G$. Since $c_n \in \mathfrak{h}(x_n)$, we have $c_n \cdot x_0 = x_n$. Hence $h_xc_n \cdot x_0 = h_x \cdot (c_n \cdot x_0) = h_x \cdot x_n$. Consequently, $\mathfrak{g}(h_x \cdot x_n) = (h_xc_n)\mathfrak{g} \wedge G \subset V_m(\mathfrak{h}_x)$ for almost every n. Hence $\mathfrak{h}(h_x \cdot x_n) \to \mathfrak{h}(x) = \mathfrak{h}(h_x \cdot x_0)$. The proof of the converse offers no difficulty, and amounts essentially to reversing the preceding argument.

24. Suppose now that Γ is a closed compact subset of G. Let $\mathfrak{H}^\Gamma$ be the totality of $\mathfrak{h}_x$'s which meet Γ. Then $\mathfrak{H}^\Gamma$ is closed and compact in $\mathfrak{H}$. For suppose $\mathfrak{h}(x_n)$ is a sequence of elements in $\mathfrak{H}^\Gamma$ such that $\mathfrak{h}(x_n) \to \mathfrak{h}(x)$. Let a_n be a point in $\mathfrak{h}(x_n) \wedge \Gamma$. Since Γ is compact, we can choose a converging subsequence $a_{n_i} \to a$, where $a \in \Gamma$. Clearly $a_{n_i}\mathfrak{g} \wedge G \to a\mathfrak{g} \wedge G$, so that $\mathfrak{h}(x_{n_i}) \to \mathfrak{h}(a \cdot x_0) \in \mathfrak{H}^\Gamma$. Since $\mathfrak{h}(a \cdot x_0) = \mathfrak{h}(x)$, $\mathfrak{H}^\Gamma$ is closed in $\mathfrak{H}$. The compactness of $\mathfrak{H}^\Gamma$ follows in similar fashion.

Let X^Γ be the totality of points x such that $\mathfrak{h}(x) \subset \mathfrak{H}^\Gamma$. Then since the correspondence $\mathfrak{h}(x) \to x$ is continuous, X^Γ is closed and compact in X_0.

Let $\Gamma_1, \Gamma_2, \cdots$ be closed compact subsets of G such that $G = \Sigma\Gamma_i$ (see §19). Then $\Sigma X^{\Gamma_i} = X_0$, and hence by (*b*), §19, at least one set in the summand, say X^{Γ_1}, must contain an inner point, say z.

Now suppose $x_n \to x_0$ in G. Then $h_z \cdot x_n \to h_z \cdot x_0 = z$. Hence almost all

the points $h_x \cdot x_n$ are in X^{Γ_1}, and the corresponding elements $\mathfrak{h}(h_x \cdot x_n)$ of $\mathfrak{H}$ are then in $\mathfrak{H}^{\Gamma_1}$. Therefore, for almost all values of n, $\mathfrak{h}(h_x \cdot x_n) \wedge \Gamma_1$ contains at least one point b_n. Since $b_n \in \Gamma_1$ and Γ_1 is closed and compact, we can choose a converging subsequence $b_{n_i} \to b$, where $b \in \Gamma_1$. Since $b_{n_i} \in \mathfrak{h}(h_x \cdot x_{n_i})$, we have $b_{n_i} \cdot x_0 = h_x \cdot x_{n_i}$. Consequently, $\mathfrak{h}(h_x \cdot x_{n_i}) = \mathfrak{h}(b_{n_i} \cdot x_0) = b_{n_i}\mathfrak{g} \wedge G \to b\mathfrak{g} \wedge G$. Since $b_{n_i} \cdot x_0 \to b \cdot x_0$ and $h_x \cdot x_{n_i} \to h_x \cdot x_0$, we have $b \cdot x_0 = h_x \cdot x_0$. Consequently, $b\mathfrak{g} \wedge G = \mathfrak{h}(h_x \cdot x_0)$, and hence $\mathfrak{h}(h_x \cdot x_{n_i}) \to \mathfrak{h}(h_x \cdot x_0)$. Hence, by the lemma (§23), $\mathfrak{h}(x_{n_i}) \to \mathfrak{h}(x_0)$. Since the correspondence $\mathfrak{h}(x) \to x$ is (1, 1) and continuous, it follows immediately that $\mathfrak{h}(x_n) \to \mathfrak{h}(x_0)$.

Suppose finally that $x_n \to x = h_x \cdot x_0$. Then $h_x^{-1} \cdot x_n \to x_0$. Hence $\mathfrak{h}(h_x^{-1} \cdot x_n) \to \mathfrak{h}(x_0)$, and hence, by the lemma, $\mathfrak{h}(x_n) \to \mathfrak{h}(h_x \cdot x_0) = \mathfrak{h}(x)$. This establishes the bi-continuity of the correspondence $\mathfrak{h}(x) \to x$.

25. Theorem. *Suppose $(G, G, G, G; x_0)$ is a transitive system. If $x = b \cdot x_0$ $(b \in G)$ and if $x_n \to x$ $(x_n \in X_0 = G \cdot x_0)$, there exists a sequence b_n such that $b_n \to b$ and $b_n \cdot x_0 = x_n$.*

Proof. Assume first that $b = e$. Then since $x_n \to x_0$, we have $\mathfrak{h}(x_n) \to \mathfrak{h}(x_0)$. Let $N(m)$ be an integer such that $N(m) < N(m + 1)$ $(m = 1, 2, \cdots)$ and such that

$$\text{(1)} \qquad \mathfrak{h}(x_n) \subset V_m(\mathfrak{h}(x_0)) = G_m\mathfrak{g} \wedge G \qquad \text{when } n > N(m) .$$

(1) implies

$$\mathfrak{h}(x_n) = c_{nm}\mathfrak{g} \wedge G, \qquad (c_{nm} \in G_m, n > N(m)) .$$

Since $c_{nm} \in \mathfrak{h}(x_n)$, we have

$$c_{nm} \cdot x_0 = x_n \qquad (n > N(m)) .$$

Now let b_n be an arbitrary point in G, when $1 \leqq n \leqq N(1)$, and let $b_n = c_{nm}$, when $N(m) < n \leqq N(m + 1)$, $m = 1, 2, \cdots$. Then clearly $b_n \to e$ and $b_n \cdot x_0 = x_n$, and the proof is complete for the case $b = e$. For the general case, we note that $b^{-1} \cdot x_n = x_0$. Hence there exist points g_n such that $g_n \to e$ and $g_n \cdot x_0 = b^{-1} \cdot x_n$. If we let $b_n = bg_n$, we have $b_n \cdot x_0 = x_n$ and $b_n \to b$.

Corollary. *If B is any open subset of G, then $B \cdot x_0$ is open.*

For let $b \cdot x_0$ $(b \in B)$ be an arbitrary point in $B \cdot x_0$. Let x_n be an arbitrary point sequence in X_0 converging to $b \cdot x_0$, and let a sequence b_n be chosen in G such that $b_n \to b$ and $b_n \cdot x_0 = x_n$. Since almost all the b_n's are in B, almost all the x_n's are in $B \cdot x_0$. Hence $b \cdot x_0$ is an inner point of $B \cdot x_0$.

The corollary states in effect that transitivity over a given region implies transitivity over any arbitrarily small portion of that region.

26. **Partial subgroups.** Let $\mathfrak{G}$ be a group and $\mathfrak{g}$ a subset of $\mathfrak{G}$ such that $\mathfrak{g}^{-1} = \mathfrak{g}$ and such that $\mathfrak{g}\mathfrak{g} \wedge G = \mathfrak{g} \wedge G$, for some suitably chosen neighborhood G of e. We shall call $\mathfrak{g}$ a *partial subgroup of* $\mathfrak{G}$. The set $[\mathfrak{g}] = \mathfrak{g} + \mathfrak{g}\mathfrak{g} + \cdots$

is obviously a subgroup of $\mathfrak{G}$. For $[\mathfrak{g}]\,[\mathfrak{g}] = [\mathfrak{g}]$ and $[\mathfrak{g}]^{-1} = \mathfrak{g}^{-1} + \mathfrak{g}^{-1}\mathfrak{g}^{-1} + \cdots = \mathfrak{g} + \mathfrak{g}\mathfrak{g} + \cdots = [\mathfrak{g}]$.

Suppose that $\mathfrak{g}$ is a partial subgroup of $\mathfrak{G}$, and that there is a neighborhood A of e such that $\mathfrak{g} \wedge A$ is closed in A, while $[\mathfrak{g}]$ is not closed in $\mathfrak{G}$. We shall say that $\mathfrak{g}$ is *recurrent*. For example, let $\mathfrak{G}$ be the 2-dimensional toroidal group; its elements are then number pairs $a = (a_1, a_2)$, where $a = b$ if $a_1 = b_1$, $a_2 = b_2 \pmod{2\pi}$. Group multiplication is given by $(ab)_i = a_i + b_i \pmod{2\pi}$ $(i = 1, 2)$. Let θ be incommensurable with 2π, and let $\mathfrak{g}$ consist of the pairs $(a, \theta a)$ $(-1 < a < 1)$. Then $\mathfrak{g}$ is a partial subgroup of $\mathfrak{G}$ and is recurrent, since $[\mathfrak{g}]$ fills $\mathfrak{G}$ densely but is closed in no region in $\mathfrak{G}$. If we drop the modulus 2π, $\mathfrak{G}$ becomes a vector group and is simply connected, but $\mathfrak{g}$ is no longer recurrent. We do not know whether or not simply connected groups, even if they are Lie groups, can contain recurrent partial subgroups.

27. Lemma. *Let $\mathfrak{g}$ be a partial subgroup in the group $\mathfrak{G}$, let $G_1, G_2, \cdots$ be neighborhoods closing down on e, and let $\mathfrak{g}_k = G_k \wedge \mathfrak{g}$ $(k = 1, 2, \cdots)$. If g is a point of $[\mathfrak{g}]$ and n a positive integer, there exists an M $[= M(n)]$ such that $\mathfrak{g}_m g \subseteq g\mathfrak{g}_n$, when $m > M$.*

Proof. Suppose that $g = g_1 g_2 \cdots g_p$ $(g_i \in \mathfrak{g})$. If our assertion is false, we can choose an increasing sequence $m_1, m_2, \cdots$ and a point $h_{m_i} \in \mathfrak{g}_{m_i}$ $(i = 1, 2, \cdots)$ such that $h_{m_i} g \bar{\in} g\mathfrak{g}_n$. Now let

$$b_{m_i} = g^{-1} h_{m_i} g = g_p^{-1} \cdots g_2^{-1} g^{-1} h_{m_i} g_1 g_2 \cdots g_p .$$

Then we may write

$$b_{m_i} = g_p^{-1} \cdots g_2^{-1} h'_{m_i} g_2 \cdots g_p, \quad h' = g_1^{-1} h_{m_i} g_1 .$$

Since $h_{m_i} \to e$, it follows that for almost all m_i's, $h_{m_i} g_1 \in G$, where G is the open set which occurs in the definition of partial subgroup. Since $h_{m_i} \in \mathfrak{g}$ and $g_1 \in \mathfrak{g}$, it follows that $h_{m_i} g_1 \in \mathfrak{g}\mathfrak{g}$. Since $G \wedge \mathfrak{g}\mathfrak{g} = G \wedge \mathfrak{g}$, it follows that for almost all m_i's, $g_1^{-1} h_{m_i} g_1 \in \mathfrak{g}\mathfrak{g}$. Since $g_1^{-1} \in \mathfrak{g}^{-1} = \mathfrak{g}$ and since, for almost all m_i's, $g_1^{-1}(h_{m_i} g_1) \in G$ (because $h_{m_i} \to e$), it follows similarly that for almost all m's, $h'_{m_i} = g_1^{-1}(h_{m_i} g_1) \in \mathfrak{g}$. Hence almost all the h'_{m_i}'s are in $\mathfrak{g}_n$. By exactly the same argument, we can show that for almost all m_i's, $g_2^{-1} h'_{m_i} g_2 \in \mathfrak{g}_n$, and on repeating this we finally obtain the result that almost all the b_{m_i}'s are in $\mathfrak{g}_n$. Since $h_{m_i} = g b_{m_i}$, it follows that almost all the h_{m_i}'s are in $g\mathfrak{g}_n$. This is a contradiction.

28. We now recall that $\mathfrak{G}$ is separable (§19). If $\mathfrak{g}$ is a partial subgroup, there exists a denumerable set $\Delta \subseteq \mathfrak{g}$ which is everywhere dense in $\mathfrak{g}$. We assert that *every set $c\mathfrak{g}_n$, where $c \in [\mathfrak{g}]$ and $\mathfrak{g}_n = G_n \wedge \mathfrak{g}$, contains points of the denumerable set $[\Delta] = \Delta + \Delta\Delta + \cdots$.* To prove this, we shall first introduce a new topology into $[\mathfrak{g}]$. The neighborhoods of an arbitrary point g in $[\mathfrak{g}]$ are to be the sets of $g\mathfrak{g}_k$ $(k = 1, 2, \cdots)$. The Hausdorff axioms can be immediately verified. Moreover, products ab remain continuous in the new $[\mathfrak{g}]$-

topology. To see this, let $ab\mathfrak{g}_n$ be a neighborhood of ab. Choose N so that $\mathfrak{g}_p\mathfrak{g}_q \subseteq \mathfrak{g}_n$ when $p > N$, $q > N$. Let k be an integer $> N$. Choose (by the lemma, §27) an integer M such that $\mathfrak{g}_j b \subseteq b\mathfrak{g}_k$, when $j > M$. Now let h be an integer $> \max(j, N)$. Then $\mathfrak{g}_h b \subseteq b\mathfrak{g}_k$, hence $\mathfrak{g}_h b\mathfrak{g}_h \subseteq b\mathfrak{g}_k\mathfrak{g}_k \subseteq b\mathfrak{g}_n$, and $a\mathfrak{g}_h b\mathfrak{g}_h \subseteq ab\mathfrak{g}_n$. If $a_q \to a$ and $b_q \to b$ relative to the new $[\mathfrak{g}]$-topology, almost all the a's are in $a\mathfrak{g}_h$, and almost all the b's are in $b\mathfrak{g}_h$. Almost all the points $a_q b_q$ are in $a\mathfrak{g}_h b\mathfrak{g}_h \subseteq ab\mathfrak{g}_n$. Hence $a_q b_q \to ab$. Our assertion follows immediately. For suppose that $c = g_1 g_2 \cdots g_t$ $(g_i \in \mathfrak{g})$. There are points $d_1^{(q)}, \cdots, d_t^{(q)}$ in Δ such that $d_i^{(q)} \to g_i$ $(i = 1, \cdots, t)$. Hence in the new $[\mathfrak{g}]$-topology

$$d^{(q)} = d_1^{(q)} d_2^{(q)} \cdots d_t^{(q)} \to g .$$

Hence an arbitrary neighborhood $c\mathfrak{g}_n$ of c contains almost all the $d_i^{(q)}$'s, and since $d^{(q)} \subseteq \Delta$, our assertion is proved.

29. Let $\mathfrak{G}$ still be a group, and let $(\mathfrak{G}, \mathfrak{X})$ be a realization, and $(G, G, G, G; x_0)$ a transitive system in $(\mathfrak{G}, \mathfrak{X})$. The set $\mathfrak{g}$ defined in §20 is obviously a partial subgroup on account of (1), §20, and we shall now show:

If $\mathfrak{g}$ is non-recurrent, there exists a neighborhood H of e such that $H \wedge \mathfrak{g} = H \wedge [\mathfrak{g}]$.[17]

Proof. We assert first that there can be chosen a denumerable set of points $\{x_n\}$ in $X_0 = G \cdot x_0$ such that (see §23)

$$G \wedge [\mathfrak{g}] = \mathfrak{h}(x_1) + \mathfrak{h}(x_2) + \cdots .$$

For let Δ $(\Delta \subseteq \mathfrak{g})$ be a denumerable set which is dense in $\mathfrak{g}$. Let

$$L = [\Delta] = \{l_1, l_2, \cdots\}, \qquad K = \mathfrak{h}(l_1 \cdot x_0) + \mathfrak{h}(l_2 \cdot x_0) + \cdots .$$

Since $l_i \in [\mathfrak{g}]$, we have $l_i\mathfrak{g} \subseteq [\mathfrak{g}]$ $(i = 1, 2, \cdots)$. Consequently, $l_i\mathfrak{g} \wedge G \subseteq [\mathfrak{g}] \wedge G$; that is, $\mathfrak{h}(l_i \cdot x_0) \subseteq [\mathfrak{g}] \wedge G$, and hence $K \subseteq [\mathfrak{g}] \wedge G$. On the other hand, let f be a point of $[\mathfrak{g}] \wedge G$ and let n be chosen so large that $f\mathfrak{g}_n \subseteq G$, where $\mathfrak{g}_n = G_n \wedge \mathfrak{g}$ (§27). Now by §28, $f\mathfrak{g}_n$ contains points of $[\Delta]$, and we may suppose that one of them is l_1. Since $l_1 \in f\mathfrak{g}_n \subseteq f\mathfrak{g}_n \wedge G \subseteq f\mathfrak{g} \wedge G$, we have $l_1 \cdot x_0 = f \cdot x_0$, and consequently $f\mathfrak{g}_n \wedge G = \mathfrak{h}(l_1 \cdot x_0)$, [(3), §20]. Hence $f \in \mathfrak{h}(l_1 \cdot x_0)$, $f \in \mathfrak{h}(l_1 \cdot x_0)$, $[\mathfrak{g}] \wedge G \subseteq K$, and $[\mathfrak{g}] \wedge G = K$. If we let $x_i = l_i \cdot x_0$, our assertion is proved.

Let $\chi = \{x_1, x_2, \cdots\}$. *We assert now that χ is closed in X_0.* For suppose that $x_{n_i} \to x$ $(x_{n_i} \in \chi, x \in X_0)$. Then (§23) $\mathfrak{h}(x_{n_i}) \to \mathfrak{h}(x)$, which we may also write $\mathfrak{h}(l_{n_i} \cdot x_0) \to \mathfrak{h}(l \cdot x_0)$ $(l \in G, l \cdot x_0 = x)$. By the theorem of §25, we may replace l_{n_i} by l'_{n_i}, such that $l'_{n_i} \to l$. Since $l'_{n_i} \in \mathfrak{h}(l'_{n_i} \cdot x_0) = \mathfrak{h}(l_{n_i} \cdot x_0) \subseteq [\mathfrak{g}]$, we have $l'_{n_i} \in [\mathfrak{g}] \wedge G$. Since $\mathfrak{g}$ is obviously closed in G and is non-recurrent, $[\mathfrak{g}]$ is closed in $\mathfrak{G}$. Consequently $l \in [\mathfrak{g}] \wedge G$, and $\mathfrak{g}(l \cdot x_0) \subseteq [\mathfrak{g}]$. Hence $x = l \cdot x_0 \in \chi$, and χ is closed in X_0.

[17] This theorem probably holds not merely for $\mathfrak{g}$, which requires for its definition the existence of a realization, but for every partial subgroup of $\mathfrak{G}$ which is closed in some neighborhood of e.

Suppose now that our theorem is false. If we assume, as we may, that $G_n \subsetneq G$ $(n = 1, 2, \cdots)$, G_1 will surely contain a point of $[\mathfrak{g}]$ which is not in $\mathfrak{g}$, and therefore G_1 is intersected by some $\mathfrak{h}(l_i \cdot x_0)$—we may say $\mathfrak{h}(l_{n_1} \cdot x_0)$—different from $\mathfrak{h}(x_0) = \mathfrak{g} \wedge G$. Since $\mathfrak{h}(l_{n_1} \cdot x_0)$ is closed in G, we can choose $m_2 > 1$ so large that G_{m_2} fails to meet $\mathfrak{h}(l_{n_1} \cdot x_0)$. But since the theorem is being denied, G_{m_2} is also intersected by $\mathfrak{h}(l_{n_2} \cdot x_0)$ different from $\mathfrak{h}(x_0)$. By the choice of m_2, $\mathfrak{h}(l_{n_2} \cdot x_0)$ is also different from $\mathfrak{h}(l_{n_1} \cdot x_0)$. On continuing in this manner, we obtain a sequence $\mathfrak{h}(l_{n_i} \cdot x_0)$ $(i = 1, 2, \cdots)$, and $\mathfrak{h}(l_{n_i} \cdot x_0) \wedge G_{m_i} \neq 0$ $(1 = m_1 < m_2 < \cdots)$. Let a_i be a point in $\mathfrak{h}(l_{n_i} \cdot x_0) \wedge G_{m_i}$, and let $\alpha = \{a_1, a_2, \cdots\}$. Since the sets $\mathfrak{h}(l_{n_i} \cdot x_0)$ are mutually exclusive, the a's are distinct, and since $a_n \to e$, α has e for a limit point. Clearly, then, the set $[\alpha] = \alpha + \alpha\alpha + \cdots$ is dense in itself, and so is $\alpha \wedge G_1$. This last set is denumerable, say $[\alpha] \wedge G_1 = \{b_1, b_2, \cdots\}$, where the b's are distinct. Let $\chi_b = \{b_1 \cdot x_0, b_2 \cdot x_0, \cdots\}$. Then obviously $\chi_b \subsetneq \chi$. Moreover, χ_b is dense in itself. For if $\mathfrak{h}(b_i \cdot x_0)$ is a limit point of $\{\mathfrak{h}(b_j \cdot x_0)\}$, then by §23, $b_i \cdot x_0$ is a limit point of $\{b_j \cdot x_0\} = \chi_b$.

Since $\mathfrak{G}$ is separable, we may assume the G_n's so chosen that $\bar{G}_1 \subset G$. Then $\bar{G}_1 \cdot x_0 \subsetneq G \cdot x_0$. Since $\bar{G}_1 \cdot x_0$ is closed, while $G \cdot x_0$ is open, it follows that $\bar{G}_1 \cdot x_0 \subset G \cdot x_0$. Therefore, since $\chi_b \subset G_1 \cdot x_0 \subset \bar{G}_1 \cdot x_0 \subset G \cdot x_0$, every limit point of χ_b is in $G \cdot x_0$; and since $\chi_b \subsetneq \chi \subset G \cdot x_0$, and χ is closed in $G \cdot x_0$, every limit point of χ_b is in χ. Hence $\bar{\chi}_b \subsetneq \chi$. But the set $\bar{\chi}_b$ is perfect, and therefore non-denumerable (theorem of Cantor for complete spaces), whereas χ is denumerable. This contradiction proves the theorem.

30. *If* $\mathfrak{g}$ *is non-recurrent, there can be chosen a neighborhood* $A \subsetneq G$ *of* e *such that if* $a \in A$,

$$a[\mathfrak{g}] \wedge A = \mathfrak{h}(a \cdot x_0) \wedge A \,. \tag{1}$$

In fact, we need only choose A such that $A^{-1} = A$ and $AA \subsetneq H$, where H is defined in the last theorem. Then if $a \in A$, $g \in [\mathfrak{g}]$, and $ag \in A$, $g \in a^{-1}A \subset AA \subsetneq H$, $g \in H \wedge \mathfrak{g} = \mathfrak{g}$, and $a[\mathfrak{g}] \subsetneq a\mathfrak{g} \wedge G = \mathfrak{h}(a \cdot x_0)$. Since ag is an arbitrary point of $a\mathfrak{g} \wedge A$, we have proved that $a[\mathfrak{g}] \wedge A \subsetneq \mathfrak{h}(a \cdot x_0)$. Since $\mathfrak{h}(a \cdot x_0) \subsetneq [\mathfrak{g}]$, we have $A \wedge \mathfrak{h}(a \cdot x_0) \subsetneq A \wedge [\mathfrak{g}]$, and (1) follows.

Imbedding theorems

31. We have already remarked that the classical theory of continuous transformation groups deals with partial structures and their systems; the modern theory, on the other hand,—or at least that part of it which studies properties in the large—deals with total structures, that is, (total) groups and total realizations of them. We shall now show that many structures of the first type can, in a sense, be imbedded in structures of the second type, so that in all probability the partial structures do not constitute an essentially more general class than the total ones.

DEFINITION. Let (A,B) and (A^*,B^*) be systems in the group structures $\mathfrak{G}$ and $\mathfrak{G}^*$, and suppose there exist homeomorphisms between A and A^*,B and

B^*, AB and A^*B^* such that if $a \to a^*$, $b \to b^*$, then $ab \to a^*b^*$. We shall say that (A,B) and (A^*,B^*) are *isomorphic*: $(A,B) \cong (A^*,B^*)$.

32. Let $\mathfrak{G}_r$ be a connected (total) group in which there can be introduced a coördinate system $a_1, \cdots, a_r$ extending over a neighborhood A of e, relative to which there exists a system (E,E) $(e \in E \subsetneq EE \subsetneq A)$ of class $C^{(2)}$. *$\mathfrak{G}_r$ is called an r-parameter Lie group.*[18] Cartan has proved that if a partial group $\mathfrak{G}$ possesses a system (E,E) $(e \in E)$ of class $C^{(2)}$, that portion of G^r which lies near e can be imbedded in a (total) Lie group $\mathfrak{G}_r$.[19] Stated more precisely, there exists a Lie group $\mathfrak{G}_r$ and a system (A,A) $(e \in A \subsetneq E)$ in $\mathfrak{G}^r$ which is isomorphic to a system (A^*,A^*) in $\mathfrak{G}_r$. We propose now to show that every system of degree 2 in a group structure $\mathfrak{G}^r$ can be similarly imbedded in a Lie group if it is part of a sufficiently regular system of degree 3.

THEOREM. *If (B,A,C) is a system in a group structure $\mathfrak{G}^r$ and if there exists a coördinate system in which (B,A) and (B,AC) are of class $C^{(2)}$, there exists a system (A_0,C_0), where $A_0 \subsetneq A$, $C_0 \subsetneq C$, and a Lie group $\mathfrak{G}_r$ such that (A_0,C_0) is isomorphic to a system (A^*,A^*) in $\mathfrak{G}_r$.*

Proof. We observe first that (A,C) is of class $C^{(1)}$; at least, this will be so if we replace A and C by suitable open subsets of themselves,[20] as is obviously permissible. It follows that if we write $C = X$, $c = x$, $ac = a \cdot x$, $bac = ba \cdot x$, the results of §§15, 16 are applicable. Consider the system (A_2,A_2) of §15, with the identity a_0. Since $a_0 \in A_3 \subsetneq A_2$ (A_3 as in §16), (A_3,A_3) is a system of the same type. We shall show first that there are systems (A_3', A_3') $(a_0 \in A_3' \subsetneq A_3)$ and (A_3',C_3') $(C_3' \subsetneq C_3)$ such that

$$(A_3',A_3') \cong (A_3',C_3') \, . \tag{1}$$

Consider the homeomorphisms (§16) of A_3: $a \to \nu(a) = a$, $a \to \lambda(a)$, $a \to \mu(a)$. Now let A_3' be a neighborhood of a_0 such that $A_3'A_3' \subsetneq A_3$, and let $C_3' = X_3' =$

[18] Cartan, [2], p. 15.

[19] Cartan's theorem ([2], p. 19) asserts that for every allowable set of structure constants there exists a total Lie group with these constants. An allowable set of structure constants is determined for $\mathfrak{G}^r$ when (E,E) is of class $C^{(2)}$ and the Lie group with the same constants has the same structure as $\mathfrak{G}^r$ in the neighborhood of the identity.

[20] To prove this, note first that since the transformation $b \to ba$ is (1,1), we can choose $A^0 \subseteqq A$ and $B^0 \subseteqq B$ such that $\left| \dfrac{\partial (ba)_i}{\partial a_j} \right| \neq 0$ when $b \in B^0$, $a \in A^0$. Since $\dfrac{\partial b_h}{\partial a_k} = \dfrac{\partial b_h}{\partial (ba)_j} \dfrac{\partial (ba)_j}{\partial a_k}$ (summed with respect to the repeated index) and since (B, A) is of the class $C^{(2)}$, the derivatives $\partial b_h / \partial a_k$ exist and are continuous for B^0, A^0. In the relation $(ba)c = b(ac)$, let $(ba)_i$, a_i, and c_i be taken as the independent variables. Then $\dfrac{\partial ((ba)c)_i}{\partial a_k} = 0$. Hence

$$0 = \frac{\partial (b(ac))_i}{\partial a_k} = \frac{\partial (b(ac))_i}{\partial b_h} \frac{\partial b_h}{\partial a_k} + \frac{\partial (b(ac))_i}{\partial (ac)_h} \frac{\Delta (ac)_h}{\Delta a_k} + \epsilon_i(ba, a, c, \Delta a) \, ,$$

where $\epsilon_i \to 0$ with Δa_k. By the same argument as above, $\left| \dfrac{\partial (b(ac))_i}{\partial (ac)_h} \right| \neq 0$ for, say, $a \in A^{00} \subseteqq A^0$, $c \in C^0$. Hence for these values of a and c $\lim \dfrac{\Delta (ac)_h}{\Delta a_k}$ exists and is continuous.

$\lambda(A_3')$. Let a_1, a_2 be arbitrary points in A_3'. Then $a_1a_3 \in A_3$ and $\lambda(a_2) = x_2$, where $x_2 \in X$. By the definition of $\lambda(a)$, $a_2 \cdot x_0 = a_0 \cdot x_2$ (§16). Hence $\mu(a_1a_2) = a_1a_2 \cdot x_0 = \tau_{a_1a_2}(a_0 \cdot x_0) = \tau_{a_1}\tau_{a_2}(a_0 \cdot x_0) = \tau_{a_1}(a_2 \cdot x_0) = \tau_{a_1}(a_0 \cdot x_2) = a_1 \cdot x_2 = a_1 \cdot \lambda(a_2) = \nu(a_1) \cdot \lambda(a_2)$, and $\mu(a)$ is a homeomorphism between $A_3'A_3'$ and $A_3' \cdot X_3' = A_3'C_3'$. (1) now follows from the relation $\mu(a_1a_2) = \nu(a_1) \cdot \lambda(a_2)$. From the theorem of Cartan, there exists a system (A_0,A_0) $(a_0 \in A_0 \subseteq A')$ and a group $\mathfrak{G}_r^*$ such that (A_0,A_0) is isomorphic to a system (A^*,A^*) in $\mathfrak{G}_r^*$. Taking $C_0 = \lambda(A_0)$, we have $(A_0,C_0) \cong (A_0,A_0) \cong (A^*,A^*)$; and the theorem is proved.

33. Let us now consider isomorphism between realization systems. We shall define only a special type of such isomorphism: the systems $(A;x_0)$ and $(A;x_0^*)$ $(A \subseteq \mathfrak{G}, x_0 \in \mathfrak{X}, x_0^* \in \mathfrak{X}^*)$ are *isomorphic* provided there exists a homeomorphism between $A \cdot x_0$ and $A \cdot x_0^*$ under which $a \cdot x_0$ and $a \cdot x_0^*$ (for every a in A) are corresponding points.

Theorem. *Let $(\mathfrak{G},\mathfrak{X})$ be a realization of a group with a transitive system $(B \cdot x_0)$. If $\mathfrak{G}$ contains no recurrent partial subgroups, there exists a transitive total realization $(\mathfrak{G},\mathfrak{X}^*)$, a point x_0^* and an open set A $(e \in A \subseteq B)$ such that $(A \cdot x_0) \cong (A \cdot x_0^*)$.*

This theorem says in effect that the partial realization $(\mathfrak{G},\mathfrak{X})$, or at least a characteristic portion of it, can be considered as imbedded in a total realization of $\mathfrak{G}$.

Proof. Let G be a neighborhood of e such that $GGGG \subseteq B$, and $G^{-1} = G$. Then $(G,G,G,G;x_0)$ is a transitive system. Let $\mathfrak{g}$ be the partial subgroup defined in §20, and let $\bar{\mathfrak{g}} = [\mathfrak{g}]$ (§26). Let x^*, y^*, $\cdots$ be the left cosets of $\bar{\mathfrak{g}}$, and denote their totality converted into a space as in §22, by $\mathfrak{X}^*$.

Let a be an arbitrary point of $\mathfrak{G}$, and let $x^* = b\bar{\mathfrak{g}}$. Then $ab\bar{\mathfrak{g}}$ is a left coset of $\bar{\mathfrak{g}}$, and hence a point of $\mathfrak{X}^*$; denote it by $a \cdot x^*$. Suppose $a \cdot x^* = a \cdot y^*$. This would imply that if $x^* = b\bar{\mathfrak{g}}$, $y^* = c\bar{\mathfrak{g}}$, then $ab\bar{\mathfrak{g}} = ac\bar{\mathfrak{g}}$, and hence $b\bar{\mathfrak{g}} = c\bar{\mathfrak{g}}$, i.e., $x^* = y^*$. It is obvious, moreover, that $a \cdot (b \cdot x^*) = ab \cdot x^*$, and finally, the continuity of $a \cdot x^*$ is a matter of immediate verification. Hence we have a total realization $(\mathfrak{G},\mathfrak{X}^*)$, which is obviously transitive.

Let $\mathfrak{H}$ be the space constructed by the aid of $\mathfrak{g}$, as in §21. Since $\mathfrak{g}$ is non-recurrent, there exists by §30 an open set $A \subseteq G$ containing e and such that

$$a\bar{\mathfrak{g}} \wedge A = \mathfrak{h}(a \cdot x_0) \wedge A \;, \tag{1}$$

when $a \in A$. Let $\mathfrak{H}$ be the totality of sets $\mathfrak{h}(a \cdot x_0)$ which meet A, and let $\mathfrak{X}_A^*$ be the totality of sets x^* which meet A. A (1,1) correspondence can be established between $\mathfrak{H}_A$ and $\mathfrak{X}_A^*$ as follows: Let x^* be an arbitrary element of $\mathfrak{X}_A^*$ (it is of the form $a\bar{\mathfrak{g}}$, where $a \in A$). By (1) we have

$$x^* \wedge A = a\bar{\mathfrak{g}} \wedge A = \mathfrak{h}(a \cdot x_0) \wedge A \;. \tag{2}$$

Then $\mathfrak{h}(a \cdot x_0)$ is a set in $\mathfrak{H}_A$ and the desired correspondence will be obtained if we make $x^* = a\bar{\mathfrak{g}}$ correspond to $\mathfrak{h}(a \cdot x_0)$. We assert that the correspondence is

bi-continuous. For suppose $x_n^* \to x^*$ in $\mathfrak{X}_A^*$. Then by §25 we may write $x_n^* = a_n\mathfrak{g}$, $x^* = a\mathfrak{g}$, $a_n \to a$ in A. The corresponding elements in $\mathfrak{H}_A$ are $\mathfrak{h}(a\cdot x_0)$ and $\mathfrak{h}(a_n\cdot x_0)$, and by §23, $\mathfrak{h}(a_n\cdot x_0) \to \mathfrak{h}(a\cdot x_0)$. Conversely, let $\mathfrak{h}(b_n\cdot x_0)$, $\mathfrak{h}(b\cdot x_0)$ be sets in $\mathfrak{H}_A$. Then we can write $\mathfrak{h}(b_n\cdot x_0) = \mathfrak{h}(a_n\cdot x_0)$, $\mathfrak{h}(b\cdot x_0) = \mathfrak{h}(a\cdot x_0)$, where a, $a_n \in A$ and where, by §25, it may be assumed that $a_n \to a$. The corresponding sets in $\mathfrak{X}_A^*$ are $x_n^* = a_n\mathfrak{g}$ and $x^* = a\mathfrak{g}$, and by §22, $x_n^* \to x^*$. Hence the correspondence in question defines a homeomorphism $\mathfrak{X}_A^* \to \mathfrak{H}_A$.

Under the homeomorphism $\mathfrak{H} \to G\cdot x_0$ of §23, $\mathfrak{H}_A$ corresponds to a subset of $G\cdot x_0$, and from the definition of the correspondence, that set is obviously $A\cdot x_0$. On the other hand, it is clear that $\mathfrak{X}_A^*$ is the set $A\cdot x_0^*$, where $x_0^* = \mathfrak{g}$. Hence we have a homeomorphism $A\cdot x_0 \to A\cdot x_0^*$. Finally, let a be an arbitrary point in A. Then $a\cdot x_0$ corresponds, under the homeomorphism $A\cdot x_0 \to \mathfrak{H}_A$, to the element $\mathfrak{h}(a\cdot x_0)$, which in turn corresponds in $\mathfrak{X}_A^* = A\cdot x_0^*$ to $a\mathfrak{g}$, that is, to $a\cdot x_0^*$. Thus under the homeomorphism, $A\cdot x_0 \to A\cdot x_0^*$, $a\cdot x_0$ corresponds to $a\cdot x_0^*$, and we have $(A;x_0) \cong (A;x_0^*)$.

34. We shall now reverse the situation of the preceding theorem and assume that $\mathfrak{G}$ is partial and $(\mathfrak{G},\mathfrak{X})$ total. In this case the notion of "imbedding" depends on isomorphisms of the form $(G;\mathfrak{X}) \cong (G^*;\mathfrak{X})$. This relation means that there exists a homeomorphism between G and G^* such that[21] if $ab = c$ in G, then $a^*b^* = c^*$ in G^*, and such that if a is an arbitrary point in G, then $a\cdot x = a^*\cdot x$ for every x in $\mathfrak{X}$.

THEOREM. *Let $(\mathfrak{G},\mathfrak{X})$ be a total realization of a partial group, and assume that there exists a system $(G;\mathfrak{X})$ ($e \in G \subseteq H$ (see VI, §6)) such that a relation of the form $g_1\cdot x = g_2\cdot x$ ($g_1,g_2 \in G$) can not hold for every x unless $g_1 = g_2$. There exists a total group $\mathfrak{G}^*$, a realization $(\mathfrak{G}^*,\mathfrak{X})$, and a system $(G^*;\mathfrak{X})$ such that $(G;\mathfrak{X}) \cong (G^*;\mathfrak{X})$.*

Proof. Let $\mathfrak{A}$ be the totality of symbols of the form

$$(a_1a_2 \cdots a_p) \qquad (a_i \in G;\quad i = 1, \cdots, p;\quad p \geqq 1)\,.$$

We define two operations on such symbols. (A) If a_q, a_{q+1} are a pair of consecutive elements in a symbol and a_qa_{q+1} is defined and equal to b, replace a_qa_{q+1} by b in the symbol. (B) If a_q is an element in a symbol, if a_q', a_q'' are in G and $a_q'a_q'' = a_q$, replace a_q by $a_q'a_q''$. Two symbols, $(a_1 \cdots a_p)$ and $(b_1 \cdots b_r)$, one of which can be obtained from the other by a sequence of operations (A) and (B) will be called equivalent. This equivalence is obviously symmetric, reflexive, and transitive. Hence $\mathfrak{A}$ falls into mutually exclusive equivalence classes, and the class determined by $(a_1 \cdots a_p)$ will be denoted by $[(a_1 \cdots a_p)]$. We now define products of classes $a^* = [(a_1 \cdots a_p)]$, $b^* = [(b_1 \cdots b_q)]$ by the formula

$$a^*b^* = [(a_1 \cdots a_p, b_1 \cdots b_q)]. \tag{1}$$

[21] I.e., if a and b are in G and ab is defined and in G.

The validity of this definition depends on the obvious fact that if $(a_1 \cdots a_p)$ and $(b_1 \cdots b_q)$ are replaced by equivalent symbols, the class on the right is unchanged. Similarly we define

$$(2) \qquad a^{*-1} = [(a_p^{-1} \cdots a_1^{-1})].$$

It is now seen that with products defined by (1), the totality $\mathfrak{G}^*$ of classes is a group, the identity is $e^* = [(e)]$, and inverses are given by (2).

We now observe that if $a,c \in G$ and $a \neq c$, then $[(a)] \neq [(c)]$. For suppose $[(a)] = [(c)]$. Then (a) and (c) are equivalent, and there exists a sequence of symbols

$$(a),\ (a'a''),\ \cdots,\ (b_1b_2 \cdots b_q \cdots b_p),\ (b_1b_2 \cdots b_q'b_q'' \cdots b_p),\ \cdots,\ (c'c''),\ (c),$$

in which all the letters denote points of G, and $a'a'' = a, \cdots, b_q'b_q'' = b_q, \cdots, c'c'' = c$. Now $a\cdot x$, $a'\cdot(a''\cdot x), \cdots, b_1\cdot(b_2\cdot(\cdots b_q\cdot(\cdots (b_p\cdot x)) \cdots)), \cdots, c'\cdot(c''\cdot x)$, $c\cdot x$ are all defined, since $(\mathfrak{G}, \mathfrak{X})$ is a total realization, and they are all equal because of I_r, §11. Hence $a\cdot x = c\cdot x$ for every $x \in \mathfrak{X}$. This, by hypothesis, is impossible unless $a = c$.

It follows that the correspondence

$$(3) \qquad a \to [(a)]$$

is (1,1) between G and a subset G^* of $\mathfrak{G}^*$. Moreover, if $ab = c$ in G, then by (1)

$$(4) \qquad [(a)]\,[(b)] = [(a)\,(b)] = [(c)],$$

so that product relations are preserved under (3).

Let $G_1, G_2, \cdots (\subseteq G)$ be a sequence of neighborhoods closing down on e and $G_1^*, G_2^*, \cdots$ the corresponding subsets of $\mathfrak{G}^*$. Let $V_n(a^*) = a^*G_n^*$, and take the totality of V's to be a complete set of neighborhoods for $\mathfrak{G}^*$. It is easy to see that $\mathfrak{G}^*$ is thus converted into a continuous group and that the correspondence (3) is now a homeomorphism between G and G^*.

Now let a^* be an arbitrary element in $\mathfrak{G}^*$, say $a^* = [(a_1a_2 \cdots a_p)]$, and let $a^*\cdot x = a_1\cdot(a_2\cdot(\cdots a_p\cdot x) \cdots)$. The validity of this product definition depends on the fact that $a^*\cdot x$, as we have seen, is independent of the particular symbol $(a_1 \cdots a_p)$ chosen to represent a^*. In particular, if $a^* = [(a)]$ $(a \in G)$ so that a and a^* correspond under (3), then $a\cdot x = a^*\cdot x$. It is clear, moreover, that $a^*\cdot(b^*\cdot x) = a^*b^*\cdot x$. We have finally to establish the continuity of $a^*\cdot x$. Suppose first that $a_n^* \to e^*$, $x_n \to x$. For $n > N$, $a_n^* \in G$, and hence we may write

$$a_n^* = [(a_n)] \qquad (a_n \in G,\ n > N),$$

and since (5) is a homeomorphism which preserves products, we have $a_n \to e$. Hence $a_n^*\cdot x = a_n\cdot x \to x$. The general case follows immediately.

35. **Lie groups.** Let $\mathfrak{G}_r$ be a Lie group (§32) and let the coördinates $a^1, \cdots, a^r$ be a canonical system; in such a system, e is at the origin. Moreover, if $\mathfrak{g}$ is

a closed subgroup of $\mathfrak{G}_h$, it follows from a theorem of Cartan ([2], p. 24) that there exists a spherical neighborhood A_0 ($\subseteq A$) of e such that $\mathfrak{g} \wedge A_0$ is a flat ρ-cell ($0 \leqq \rho \leqq r$). We may assume that the flat ρ-space which contains $\mathfrak{g} \wedge A$ is $a^{\rho+1} = \cdots = a^r = 0$. Let $\mathfrak{k}$ be the flat space orthogonal to $\mathfrak{g}$ at e, and let B_δ be a spherical neighborhood of e with radius δ. Let the projections of a point $a = (a^1, \cdots, a^r)$ on $\mathfrak{k}$ and $\mathfrak{g}$ be $a' = (0, \cdots, 0, a^{\rho+1}, \cdots, a^r)$ and $a'' = (a^1, \cdots, a^\rho, 0, \cdots, 0)$, so that (symbolically) $a = (a', a'')$. Let $\mathfrak{k}_\epsilon = \mathfrak{k} \wedge B_\epsilon$, $\mathfrak{g}_\delta = \mathfrak{g} \wedge B_\delta$. The totality $(\mathfrak{k}_\epsilon, \mathfrak{g}_\delta)$ of points (a', a'') which are such that $a' \epsilon \mathfrak{k}$, $a'' \epsilon \mathfrak{g}$ is homeomorphic to the topological product $\mathfrak{k}_\epsilon \times \mathfrak{g}_\delta$ and is therefore an r-cell. We assert that if ϵ and δ are small enough, the single-valued correspondence $(a', a'') \to a'a''$ ($a' \epsilon \mathfrak{k}_\epsilon$, $a'' \epsilon \mathfrak{g}_\delta$) is homeomorphic between $(\mathfrak{k}_\epsilon, \mathfrak{g}_\delta)$ and $\mathfrak{k}_\epsilon\mathfrak{g}_\delta$. To see this, let $c = (c^1, \cdots, c^r) = a'a''$. The c's are functions of $a^1, \cdots, a^r$, and on account of the continuity of their derivatives, we have only to show that

$$\left|\frac{\partial c^i}{\partial a^j}\right|_{a^1 = \cdots = a^r = 0} \neq 0 .$$

In a canonical parameter system, the group product of two elements in an infinitesimal neighborhood of e is obtained simply by adding coördinates. Hence $c^i = a^i + \epsilon^i (a^1, \cdots, a^r)$ when $\lim\limits_{\tau \to 0} \epsilon^i/\tau = 0$, τ being, say, the distance from a to the origin. It follows immediately that the value of the determinant is 1; and this proves our assertion. The set $\mathfrak{k}_\epsilon\mathfrak{g}_\delta$ is therefore an r-cell, and hence, by the invariance of regionality, an open set in $\mathfrak{G}_r$.

It follows further that if a' is an arbitrary point of $\mathfrak{k}_\epsilon$, then $a'\mathfrak{g}_\delta$ has only the point a' in common with $\mathfrak{k}_\epsilon$. Now the set $\mathfrak{g}-\mathfrak{g}_\delta$ is at a distance $\geqq \delta$ from e, and hence, if ϵ is sufficiently small, $a'(\mathfrak{g} - \mathfrak{g}_\delta)$ will be near enough to $\mathfrak{g} - \mathfrak{g}_\delta$ to avoid meeting $\mathfrak{k}_\epsilon$. Hence for such an ϵ, a' meets $\mathfrak{k}_\epsilon$ only at a'. Consequently the correspondence $a' \to a'\mathfrak{g}$ between points and cosets is (1,1) when $a' \epsilon \mathfrak{k}_\epsilon$. If the cosets $a'\mathfrak{g}$ are regarded as forming a subset X_ϵ of $\mathfrak{H}$ (§22), the correspondence is a homeomorphism. Its continuity, in fact, is readily seen. As for that of its inverse, suppose that $b_n'\mathfrak{g} \to b'\mathfrak{g}$ (b_n', $b' \epsilon \mathfrak{k}_\epsilon$). Then, by §25, choose a_n^* such that $a_n^* \to b'$ and $a_n^*\mathfrak{g} = b'\mathfrak{g}$. Now $b' = b'e$ is contained in the open set $\mathfrak{k}_\epsilon\mathfrak{g}_\delta$, and hence almost all the a_n^*'s are of the form $a_n'a_n''$, where $a_n' \epsilon \mathfrak{k}_\epsilon$, $a'' \epsilon \mathfrak{g}_\delta$. Since $a_n'a_n'' \to b'$, we have $(a_n', a_n'') \to b'$. Hence $a_n' \to b'$, $a_n'' \to 0$. Since $b_n'\mathfrak{g} = a_n'a_n''\mathfrak{g} = a'\mathfrak{g}$, we must have $b_n' = a_n'$, so that $b_n' \to b'$. *Let ϵ and δ now be fixed so small that the situations just described hold.*

36. Let $\eta < \min(\epsilon, \delta)$, and let L_η be the r-cell $\mathfrak{k}_\eta\mathfrak{g}_\eta$. There can be chosen ([10], p. 19) a sequence $a_{\eta 2}, a_{\eta 2}, \cdots$ such that $\mathfrak{G}_r = \Sigma_i\, a_{\eta i}L_\eta$. The sets

$$a_{\tau i}L_\tau \qquad (i = 1, 2, \cdots ; \tau = \eta, \eta/2, \eta/3, \cdots)$$

are r-cells and obviously form a denumerable complete set of neighborhoods for $\mathfrak{G}_r$. Hence $\mathfrak{G}_r$ is a topological manifold. Since $\mathfrak{g}$ is closed, it falls into a discrete set of connected homeomorphic pieces, and the piece which contains e

is a Lie group ([2], pp. 22–24). Hence $\mathfrak{g}$ *consists of a discrete set of topological manifolds.*

The realization $(\mathfrak{G}_r, \mathfrak{H})$ defined by $a \cdot b\mathfrak{g} = ab\mathfrak{g}$ (cf. §33) is transitive, and hence $\mathfrak{H}$ is homogeneous. Now since the sets L_τ are open subsets of $\mathfrak{G}_r$, the sets $L_\tau \cdot \mathfrak{g}$ are open subsets of $\mathfrak{H}$ (§25, corollary). Since $L_\tau \cdot \mathfrak{g} = \mathfrak{k}_\tau \mathfrak{g}_\tau \cdot \mathfrak{g} = \mathfrak{k}_\tau \mathfrak{g}_\tau \mathfrak{g} = \mathfrak{k}_\tau \mathfrak{g}$, and since, by the preceding section the subset $\mathfrak{k}_\tau \mathfrak{g}$ of $\mathfrak{H}$ is homeomorphic to $\mathfrak{k}_\tau$, i.e., to an $(r - \rho)$-cell, the sets

$$a_{\tau i} L_\tau \cdot \mathfrak{g} \qquad (i = 1, 2, \cdots; \tau = \eta, \eta/2, \cdots)$$

clearly form a denumerable complete set of neighborhoods for $\mathfrak{H}$. Since $\mathfrak{G}_r$ is connected, $\mathfrak{H}$ is obviously so. *Hence $\mathfrak{H}$ is a topological $(r - \rho)$-manifold.*

Suppose in particular that $(\mathfrak{G}_r, \mathfrak{X})$ is a transitive total realization of $\mathfrak{G}_r$ and that $\mathfrak{g}$ consists of all the points a such that $a \cdot x_0 = x_0$. *Then $\mathfrak{X}$, being homeomorphic to $\mathfrak{H}$ (§23), is a topological $(r - \rho)$-manifold* (cf. [2], p. 14).

37. Theorem. *Let $\mathfrak{G}$ be a Lie group and $(\mathfrak{G}, \mathfrak{X})$ a transitive total realization, $\mathfrak{X}$ being connected. If a transformation $x \to b \cdot x$ leaves invariant the points of an open set in $\mathfrak{X}$, it leaves invariant all points of $\mathfrak{X}$.*

Proof. Suppose X is an open set such that $b \cdot x = x$ for every $x \in X$, or symbolically, $b \cdot X \equiv X$. Let $\mathfrak{g}$ be the totality of points a such that $a \cdot X \equiv X$. Let x_0 be a point in X and let $G_1, G_2 \cdots$ be connected neighborhoods closing down on e. Choose n so large that $G_n \cdot x_0 \subseteq X$, and choose m so large that $G_m^{-1} G_m \subseteq G_n$. Let g and h be points in G_m and G_n respectively. Then

$$(1) \qquad g\mathfrak{g}g^{-1} \cdot (gh \cdot x_0) = g\mathfrak{g}(h \cdot x_0) = gh \cdot x_0,$$

since $h \cdot x_0 \in X$. Now for each g in G_m, the totality of points gh—that is, the set gG_n—contains G_m; for since $g^{-1}G_m \subseteq G_n$, we have $gG_n \supseteqq G_m$. Hence for each g we have from (1) $g\mathfrak{g}g^{-1} \cdot x = x$ for every x in $G_m \cdot x_0$. This relation holds à fortiori if $g \in G_k$ $(k \geqq m)$.

Let

$$\mathfrak{g}_k = \{g\mathfrak{g}g^{-1}, g \in G_k\}, \ \mathfrak{h}_k = \mathfrak{g}_k + \mathfrak{g}_k\mathfrak{g}_k + \cdots \qquad (k \geqq m).$$

We have

$$(2) \qquad b \in \mathfrak{g} \subseteq \mathfrak{g}_k \subseteq \mathfrak{h}_k \subseteq \bar{\mathfrak{h}}_k \qquad (k \geqq m),$$

$$(3) \qquad \bar{\mathfrak{h}}_k \cdot x_0 = x \quad \text{for} \quad x \in G_m \cdot x_0 \qquad (k \geqq m),$$

$$(4) \qquad \bar{\mathfrak{h}}_m \supseteqq \bar{\mathfrak{h}}_{m+1} \supseteqq \cdots .$$

Each $\mathfrak{h}_k$ is a closed subgroup of $\mathfrak{G}$, and is therefore a discrete equi-dimensional set of mutually exclusive connected pieces, and each is a topological manifold (§36). Let $\bar{\mathfrak{h}}_k^0$ and $\bar{\mathfrak{h}}_k^*$ be the connected parts of $\bar{\mathfrak{h}}_k$ which contain e and b. If h is an arbitrary point of $\bar{\mathfrak{h}}_k^*$, we have $\bar{\mathfrak{h}}_k^* = h\mathfrak{h}_k^0$.

On account of (4), there exists an integer $q \geqq m$ such that

$$\dim \bar{\mathfrak{h}}_q = \dim \bar{\mathfrak{h}}_{q+1} = \cdots . \tag{5}$$

Let p ($> q$) be so large that $G_pG_p \subseteq G_q$. We assert that if $c \in G_p$, then

$$c\bar{\mathfrak{h}}_p c^{-1} \subseteq \bar{\mathfrak{h}}_q. \tag{6}$$

For let cjc^{-1} ($j \in \mathfrak{h}_p$) be an arbitrary point in $c\mathfrak{h}_p c^{-1}$. Then j is a product of the form

$$\prod_{i=1}^{t} g_i \mathfrak{g} g_i^{-1} \qquad (g_i \in G_p;\ i = 1, \cdots, t),$$

and hence

$$cjc^{-1} = \prod_{i=1}^{t} (c_i g_i)\mathfrak{g}(c_i g_i)^{-1}.$$

Since $c_i g_i \in G_q$ ($i = 1, \cdots, t$), we have $cjc^{-1} \in \mathfrak{h}_q$. Hence $c\mathfrak{h}_p c^{-1} \subseteq \mathfrak{h}_q$, and this relation obviously holds when we pass to the closures.

Obviously $c\mathfrak{h}_p^0 c^{-1}$ is that connected part of $c\bar{\mathfrak{h}}_p c^{-1}$ which contains e, and since it is homeomorphic to $\bar{\mathfrak{h}}_p^0$, it is of same dimension as $\bar{\mathfrak{h}}_p^0$, and by (5) of same dimension as $\bar{\mathfrak{h}}_q^0$. By (6) it follows that $c\bar{\mathfrak{h}}_p^0 c^{-1}$ is a subgroup of $\bar{\mathfrak{h}}_q^0$; but since the dimensions are equal and both are connected, it follows from a theorem of Schreier [10] that $c\bar{\mathfrak{h}}_p^0 c^{-1} = \bar{\mathfrak{h}}_q^0$. In the same way it follows from (4) and (5) that $\bar{\mathfrak{h}}_p^0 = \bar{\mathfrak{h}}_q^0$. Hence

$$c\bar{\mathfrak{h}}_0^p c^{-1} = \bar{\mathfrak{h}}_p^0. \tag{7}$$

Now $c\bar{\mathfrak{h}}_p^* c^{-1} \subseteq \bar{\mathfrak{h}}_p$ on account of (6). Since $c \in G_p$ and G_p is connected, we can join e to c by a path $c(t)$, ($c(0) = e$, $c(1) = c$) in G_p. The successive sets $c(t)\bar{\mathfrak{h}}_p^* c(t)^{-1}$ constitute a deformation of $\bar{\mathfrak{h}}_p^*$; but they are all contained in $\bar{\mathfrak{h}}_p$, and hence in the connected piece $\bar{\mathfrak{h}}_p^*$. Hence

$$c\bar{\mathfrak{h}}_p^* c^{-1} \subseteq \bar{\mathfrak{h}}_p^*. \tag{8}$$

Now $c\bar{\mathfrak{h}}_p^* c^{-1}$ is clearly one of the connected parts of $c\bar{\mathfrak{h}}_p c^{-1}$. Hence if d is a point in $c\bar{\mathfrak{h}}_p^* c^{-1}$, we have $c\bar{\mathfrak{h}}_p^* c^{-1} = d(c\mathfrak{h}_p^0 c^{-1})$. By (8), d is also in $\bar{\mathfrak{h}}_p^*$, and we have $\bar{\mathfrak{h}}_p^* = d\bar{\mathfrak{h}}_p^0$. Hence by (7)

$$c\bar{\mathfrak{h}}_p^* c^{-1} = \bar{\mathfrak{h}}_p^*. \tag{9}$$

Let X_p consist of the points x such that $\bar{\mathfrak{h}}_p^* \cdot x = x$. On account of (3), $X_p \neq 0$. Obviously X_p is closed; it is also open. For, if $y \in X_p$, then $\bar{\mathfrak{h}}_p^*(c \cdot y) = c\bar{\mathfrak{h}}_p^* c^{-1} c \cdot y = c\bar{\mathfrak{h}}_p^* \cdot y = c \cdot y$, by use of (9). Since c is an arbitrary point in G_p, $c \cdot y$ is an arbitrary point in $G_p \cdot y$. Hence $\bar{\mathfrak{h}}^* \cdot x = x$, for every $x \in G_p \cdot y$, and $G_p \cdot y \subseteq X_p$. Since $(\mathfrak{G}, \mathfrak{X})$ is transitive, $G_p \cdot y$ is open by §25. Hence X_p is open. Since $\mathfrak{X}$ is connected, it follows that $X_p = \mathfrak{X}$, and since $b \in \bar{\mathfrak{h}}_p^*$ we have $b \cdot \mathfrak{X} \equiv \mathfrak{X}$.

38. **A remark about essential parameters.** We have asserted (§17) that the assumption II′ is a consequence of the assumption that the "parameters be essential". We shall now prove that this holds, in a sense which will become obvious, when $(\mathfrak{G}_r, \mathfrak{X})$ is a transitive total realization of a Lie group. Suppose on the contrary that II′ is false. There exists a point a in $\mathfrak{G}_r$ and sequence $a_n \to a$, $b_n \to a$ $(a_n \neq b_n)$ such that $a_n \doteq b_{nn}$ (§17). Let $c_n = b_n^{-1}a_n$. Obviously $c_n \doteq e$, and $c_n \to e$. From the preceding theorem, the relation $c_n \doteq e$ implies that

$$c_n\mathfrak{X} \equiv \mathfrak{X} \qquad (n = 1, 2, \cdots).$$

Let $\mathfrak{g}$ consist of all the points a such that $a \cdot \mathfrak{X} \equiv \mathfrak{X}$. Then $\mathfrak{g}$ is a closed subgroup of $\mathfrak{G}$. We assert that dim $\mathfrak{g} > 0$. Let $c = \{c_1, c_2, \cdots\}$. Since $a_n \neq b_n$, the set c is infinite. Let $\bar{c}$ be the closure of the set $c + cc + \cdots$. Since $c_n \to e$, $\bar{c}$ contains at least one 1-parameter subgroup[22] of $\mathfrak{G}_r$, and since $\bar{c} \subseteq \mathfrak{g}$, dim $\mathfrak{g} \geqq 1$. $\mathfrak{g}$ is obviously invariant, and the group $\mathfrak{G}/\mathfrak{g}$ may be converted into a space by identifying it with the space $\mathfrak{H}$ (§22) formed with the cosets of $\mathfrak{g}$. Since $\mathfrak{g}$ is closed, the results of §35 are applicable; hence $\mathfrak{H}$ is a topological manifold, and we have dim $\mathfrak{H} = r -$ dim $\mathfrak{g}$, so that dim $\mathfrak{H} <$ dim $\mathfrak{G}_r$. On writing $a\mathfrak{g} \cdot x = a \cdot x$ we obtain a representation $(\mathfrak{H}, \mathfrak{X})$. The correspondence $a \to a\mathfrak{g}$ between $\mathfrak{G}_r$ and $\mathfrak{H}$ is clearly single-valued and continuous. Hence on introducing cartesian coördinate systems in neighborhoods of the identities of $\mathfrak{G}_r$ and $\mathfrak{H}$, the correspondence $a \to a\mathfrak{g}$ assumes the form

$$a \to a^* \qquad [a = (a_1, \cdots, a_r),\ a^* = (a_1^*, \cdots, a_\rho^*),\ \rho < r],$$

so that the a^*'s are single-valued continuous functions of the a's. Since the relation $a \cdot x = a^* \cdot x$ holds identically in x, the number of parameters in the functions $a \cdot x$ can be reduced; that is, the parameters in $(\mathfrak{G}_r, \mathfrak{X})$ are not essential. This contradiction proves the theorem.

The fundamental groups

39. Let $(\mathfrak{G}, \mathfrak{X})$ be a transitive total realization of a Lie group. In this section we obtain a relation between the ranks of certain fundamental (Poincaré) groups which occur in connection with $(\mathfrak{G}, \mathfrak{X})$.

Preliminary lemmas. Let $[a, b, c, \cdots, d]$ denote a path in $\mathfrak{G}$ beginning at a, passing successively through $b, c, \cdots$ and ending at d. Let $[y,z]$ be a path in $\mathfrak{X}$ beginning at y and ending at z. A relation of the form $[a, b, \cdots, d] \cdot x = [y, z]$ will mean that while a point a' traces the first path from a to d, the point $a' \cdot x$ will trace the second, from y to z. In case $y = z$, $a' \cdot x$ is to trace the closed path $[y, y]$ in the positive direction.

Lemma 1. *Let ξ be a closed path in $\mathfrak{X}$ defined by the function $\xi(t)$ of period* 1.

[22] This will be obvious on referring to Cartan, [2], p. 23.

There exists a $\delta > 0$ such that for each t there is an arc $[e,a]$ in $\mathfrak{G}$ with the property that

$$[e,a]\cdot\xi(t) = [\xi(t),\ \xi(t+\delta)],$$

the path on the right being the subpath of ξ traced when the parameter is increased continuously from t to $t + \delta$.

Proof. Suppose the lemma is false. There exists a sequence $t_1, t_2, \cdots$ such that a relation of the form $[e,a]\cdot\xi(t_n) = [\xi(t), \xi(t_n+1/n)]$ fails to exist for $n = 1, 2, \cdots$. On passing to a subsequence, if necessary, we may assume that $t_n \to \bar{t}$ (mod 1). Let $x_0 = \xi(\bar{t})$, let $\mathfrak{g}$ be the subgroup of $\mathfrak{G}$ corresponding to x_0, and form $\mathfrak{H}$ with the cosets of $\mathfrak{g}$ (§22). Let X_ϵ and $\mathfrak{k}_\epsilon$ be defined as in §35. X_ϵ is an open set containing x_0. Hence $\lambda > 0$ can be chosen so small that the path

$$[\xi(\bar{t}-\lambda),\ \xi(\bar{t}+\lambda)] \tag{1}$$

is contained in X_ϵ. Since X_ϵ is homeomorphic to $\mathfrak{k}_\epsilon$, each point $\xi(t)$ of this path corresponds to a point $k(t)$ in $\mathfrak{k}_\epsilon$. The correspondence is such (§35) that $k(\bar{t}) = e$, and $k(t)\cdot x_0 = \xi(t)$ $(\bar{t} - \lambda \leqq t \leqq \bar{t} + \lambda)$. The second relation can be written in the form

$$[k(\bar{t}-\lambda),\ k(\bar{t}+\lambda)]\cdot\xi(\bar{t}) = [\xi(\bar{t}-\lambda),\ \xi(\bar{t}+\lambda)]. \tag{2}$$

For m sufficiently large, the path $[\xi(t_m), \xi(t_m + 1/m)]$ is a subpath of (1). Hence there exists a subpath of the path in the left of (2), say $[a', a'']$, such that $[a', a'']\cdot\xi(t) = [\xi(t_m), \xi(t_m + 1/m)]$. In particular, $a'\cdot\xi(\bar{t}) = \xi(t_m)$. Hence

$$[a', a'']\cdot((a')^{-1}\xi(t_m)) = [\xi(t_m),\ \xi(t_m + 1/m)]. \tag{3}$$

The path $[a', a''][(a')^{-1}]$—that is, the path traced by $b(a')^{-1}$ as b traces $[a', a'']$—is of the form $[e, a]$, and by (3), $[e, a]\cdot\xi(t_m) = [\xi(t_m), \xi(t_m + 1/m)]$. This is a contradiction.

From now on let x_0 be an arbitrarily chosen but fixed point in $\mathfrak{X}$, so that the sets denoted by the symbols $\mathfrak{g}$, $\mathfrak{k}_\epsilon$, $\mathfrak{H}$, etc., defined relative to x_0 (§35) are fixed from now on.

LEMMA 2. *Let $\xi = [x_0, x_0]$ be a path in $\mathfrak{X}$ beginning and ending at x_0. There is a path $[e, a]$ in $\mathfrak{G}$ such that $[e, a]\cdot x_0 = [x_0, x_0]$.*

Proof. Suppose ξ is defined by $\xi(t)$, of period 1. Choose a δ with the property described in Lemma 1, and let points $x_0 = \xi(t_0), \xi(t_1), \cdots, \xi(t_{k-1}), \xi(t_k) = x_0$, be chosen such that

$$|\,t_i - t_{i+1}\,| < \delta \text{ (mod 1)} \qquad (i = 0, \cdots, k-1).$$

By Lemma 1, there are arcs $[e, a_i]$ such that

$$[e, a_1]\cdot x_0 = [\xi(t_0), \xi(t_1)],\ [e, a_2]\cdot\xi(t_1) = [\xi(t_1), \xi(t_2)], \text{ etc.} \tag{4}$$

Since $\xi(t_1) = a_1\cdot x_0$, $\xi(t_2) = a_2a_1\cdot x_0, \cdots$, the left members of (4) can be replaced by $[e, a_1]\cdot x_0$, $([e, a_2]a_1)\cdot x_0$, $([e, a_3]a_2a_1)\cdot x_0, \cdots$. The symbol $[e, a_1] + [e, a_2]a_1 +$

$[e, a_3]a_2a_1 + \cdots + [e, a_k]a_{k-1} \cdots a_1$ evidently represents a path, since each path in the sum begins where the preceding one ends; it is of the form $[e, a]$, and because of the relations (4), $[e, a]\cdot x_0 = [x_0, x_0]$.

LEMMA 3. *Suppose $\alpha(t)$ and $\beta(t)$ $(\bar{t} \leqq t \leqq t')$ define paths in $\mathfrak{G}$ such that*

$$\alpha(t)\cdot x_0 = \beta(t)\cdot x_0, \qquad \alpha(\bar{t}) = \beta(\bar{t}). \tag{5}$$

Either path can be deformed to the other in such a way that its initial point remains fixed while its final point remains in a fixed coset of $\mathfrak{g}$.

Proof. We may obviously assume that $\bar{t} = 0$, $t' = 1$. Because of (5), the point $\alpha(t)$, for each t, lies in the same coset of $\mathfrak{g}$ as does $\beta(t)$. We may write $\beta(t) = \alpha(t)\,\gamma(t)$ $(0 \leqq t \leqq 1)$, where $\gamma(t)$ defines a path in $\mathfrak{g}$. Let

$$\beta(t,s) = \begin{cases} \alpha(t) & (0 \leqq t \leqq s \leqq 1), \\ \alpha\left(\dfrac{t-s}{1-s}\right)\gamma(t-s) & (0 \leqq s \leqq t \leqq 1). \end{cases}$$

Observe that $\beta(t, 0) = \beta(t)$ and $\beta(t, 1) = \alpha(t)$, while, for each intermediate s, $\beta(t, s)$ is a path beginning at $\alpha(0)$ and ending at the point $\alpha(1)\gamma(1 - s)$ on the coset $\alpha(1)\mathfrak{g}$. $\beta(t, s)$ therefore defines the desired deformation.

LEMMA 4. *Let a set of points in $\mathfrak{X}$ be defined by $\xi(t, s)$ $(0 \leqq t, s \leqq 1)$, the function ξ being continuous in (t, s) and such that $\xi(0, s) = \xi(1, s) = \xi(t, 0) = x_0$. Let s be a fixed number $\geqq 0$ and $\leqq 1$, and let $\alpha(t)$ define a path such that*

$$\alpha(t)\cdot x_0 = \xi(t, \bar{s}) \qquad (0 \leqq t \leqq 1). \tag{6}$$

Let $\bar{t}$ be an arbitrary number such that $0 \leqq \bar{t} \leqq 1$. There exists an $h > 0$, independent of t, and a continuous function $\alpha(t, s)$ defined for

$$|\,\bar{t} - t\,| < h, \qquad |\,\bar{s} - s\,| < h \qquad (0 \leqq t \leqq 1; 0 \leqq s \leqq 1) \tag{7}$$

such that for these values $\alpha(t, \bar{s}) = \alpha(t)$, $\alpha(t, s)\cdot x_0 = \xi(t, s)$.

Proof. Let $\varphi(\bar{t}, t) = \alpha(\bar{t})^{-1}\,\alpha(t)$, $\psi(\bar{t}, t, s) = \alpha(\bar{t})^{-1}\,\xi(t, s)$. Now $\varphi(\bar{t}, t) = e$ when $t = \bar{t}$, and because of (6), $\psi(t, t, s) = x_0$ when $t = \bar{t}$ and $s = \bar{s}$. Hence by the principles of uniform continuity, there exists an h independent of t, $\bar{t}$ and s such that (see §35)

$$\begin{aligned} &\varphi(\bar{t}, t) \in \mathfrak{k}_\epsilon\mathfrak{g}_\delta, && \text{when} && |\,t - \bar{t}\,| < h, \\ &\psi(\bar{t}, t, s) \in X_\epsilon, && \text{when} && |\,t - \bar{t}\,| < h, \ |\,s - \bar{s}\,| < h. \end{aligned} \tag{8}$$

Now let $\bar{t}$ be fixed. Because of the correspondence $X_\epsilon \to \mathfrak{k}_\epsilon$ (§35), there exist points $\kappa(t, s)$ in $\mathfrak{k}_\epsilon$ [κ being continuous in (t, s)] such that

$$\kappa(t, s)\cdot x_0 = \psi(\bar{t}, t, s) = \alpha(\bar{t})^{-1}\xi(t, s). \tag{9}$$

From the first of the relations (8), there exist continuous functions $\kappa(t)$, $\gamma(t)$ such that $\kappa(t) \in \mathfrak{k}_\epsilon$, $\gamma(t) \in \mathfrak{g}_\delta$ and $\kappa(t)\gamma(t) = \varphi(\bar{t}, t) = \alpha(\bar{t})^{-1}\,\alpha(t)$. The functions $\kappa(t, s)$, $\kappa(t)$, $\gamma(t)$ are defined for values of t, s which satisfy (7). Since $\kappa(t)\cdot x_0 = \kappa(t)\gamma(t)\cdot x = \alpha(\bar{t})^{-1}\alpha(t)\cdot x_0$, and, by (9) and (6), $\kappa(t, \bar{s})\cdot x_0 = \alpha(\bar{t})^{-1}\xi(t, \bar{s}) =$

$\alpha(\bar{t})^{-1}\alpha(t)\cdot x_0$, it follows that $\kappa(t)\cdot x_0 = \kappa(t, \bar{s})\cdot x_0$. Therefore, since the points $\kappa(t)$ and $\kappa(t, \bar{s})$ are in $\mathfrak{k}_\epsilon$, we must have

$$\kappa(t) = \kappa(t, \bar{s}). \tag{10}$$

Let $\alpha(t, s) = \alpha(\bar{t})\kappa(t, s)\gamma(t)$ when s, t satisfy (7). Then $\alpha(t, s) = \alpha(\bar{t})\kappa(t, \bar{s})\gamma(t) = \alpha(\bar{t})\kappa(t)\gamma(t) = \alpha(\bar{t})(\alpha(\bar{t})^{-1}\alpha(t)) = \alpha(t)$, by use of (10) and $\alpha(t, s)\cdot x_0 = \alpha(\bar{t})\kappa(t, s)\cdot x_0 = \xi(t, s)$, when t, s satisfy (7). Hence $\alpha(t, s)$ has the desired properties.

LEMMA 5. *Let $\xi(t, s)$, $\alpha(t)$, $\bar{s}$, h be defined as in Lemma 4. For every s' such that*

$$0 \leqq s' \leqq 1, \qquad |\bar{s} - s'| < h, \tag{11}$$

there exists a path $\alpha'(t)$ satisfying $\alpha'(t)\cdot x_0 = \xi(t, s')$ deformable into $\alpha(t)$, end points remaining fixed.

Proof. Let $t_0 = 0, t_1, t_2, \cdots, t_{k-1}, t_k = 1$ be an increasing sequence such that $t_{i+1} - t_i < h$. Then by Lemma 4 there exist functions $\alpha_i(t, s)$ $(i = 0, \cdots, k - 1)$ such that

$$\alpha_i(t, s)\cdot x_0 = \xi(t, s) \qquad (t_i \leqq t \leqq t_{i+1}, |s - \bar{s}| \leqq h, 0 \leqq s \leqq 1), \tag{12}$$

$$\alpha_i(t, \bar{s}) = \alpha(t) \qquad (t_i \leqq t \leqq t_{i+1}). \tag{13}$$

Let s' be fixed, and let

$$\alpha_i(t_i, \bar{s}) = b_i, \quad \alpha_i(t_i, s') = c'_i, \quad \alpha_i(t_{i+1}, s') = c''_i, \quad (i = 0, \cdots, k - 1).$$

Note that if we define $b_k = \alpha(1)$, then

$$\alpha_i(t_{i+1}, \bar{s}) = \alpha_{i+1}(t_{i+1}, \bar{s}) = b_{i+1} \qquad (i = 1, \cdots, k - 1).$$

For definiteness, let us suppose that $s' > \bar{s}$. Then for the values of s, t such that $t_i \leqq t \leqq t_{i+1}$, $\bar{s} \leqq s \leqq s'$, the points $\alpha_i(t, s)$ constitute a singular rectangle with vertices $b_i, c'_i, c''_i, b_{i+1}$. One edge of this rectangle coincides by (13) with the subpath $[b_i, b_{i+1}]$ of $\alpha(t)$. The remaining edges constitute a second path $[b_i, c'_i, c''_i, b_{i+1}]$ joining b_i to b_{i+1}. This path can be deformed across the singular rectangle to the path $[b_i, b_{i+1}]$, end points remaining fixed. Now observe that

$$\begin{aligned} \alpha_i(t_{i+1}, s)\cdot x_0 &= \alpha_{i+1}(t_{i+1}, s)\cdot x_0 = \xi(t_{i+1}, s), \\ \alpha_i(t_{i+1}, \bar{s}) &= \alpha_{i+1}(t_{i+1}, \bar{s}) = \alpha(t_{i+1}) \qquad (\bar{s} \leqq s \leqq s'). \end{aligned} \tag{14}$$

Hence by Lemma 3, the path $[b_{i+1}, c''_i]$, defined by $\alpha_i(t_{i+1}, s)$, can be deformed to the path $[b_{i+1}, c'_{i+1}]$, defined by $\alpha_{i+1}(t_{i+1}, s)$, the initial point remaining fixed and the final point tracing a path $[c''_i, c'_{i+1}]$ in some fixed coset of $\mathfrak{g}$, so that

$$[c''_i, c'_{i+1}]\cdot x_0 = c''_i\cdot x_0 = c'_{i+1}\cdot x. \tag{15}$$

Let $[c'_i, c''_i]$ be the path defined by $\alpha(t, s')$ $(t_i \leqq t \leqq t_{i+1})$. Because of (12),

$$[c'_i, c''_i]\cdot x_0 = [\xi(t_i, s'), \xi(t_{i+1}, s')], \tag{16}$$

and hence if we join the paths $[b_0, c_0']$, $[c_0', c_0'']$, $[c_0'', c_1']$, $\cdots$, $[c_{k-1}', c_{k-1}'']$,$[c_{k-1}'', b_k]$ forming a path α', as a point a moves from b_0 to b_k along α', $a \cdot x$ moves along the path defined by $\xi(t, s')$, because of (16), pausing (by (15)) at each point $\xi(t_i, s')$ while a traces $[c_i'', c_{i+1}']$ and pausing at $\xi(0, s')$ and $\xi(1, s')$ while a traces $[b_0, c_0']$ and $[c_{k-1}'', b_k]$. Hence α' can be defined by a suitably chosen function $\alpha'(t)$ such that $\alpha'(t) \cdot x_0 = \xi(t, s')$ $(0 \leqq t \leqq 1)$. It is clear from the construction of α' that it can be deformed to $\alpha(t)$, its end points remaining fixed.

THEOREM.[23] *Let $\alpha(t)$ $(0 \leqq t \leqq 1)$ define a path in $\mathfrak{G}$ beginning at e and such that $\alpha(1) \cdot x_0 = x_0$. If the closed path in $\mathfrak{X}$, defined by $\alpha(t) \cdot x_0$, is deformable to x_0 through a family of paths beginning and ending at x_0, then $\alpha(t)$ can be deformed to a path in $\mathfrak{g}$ through a family of paths beginning at e and ending at points in $\mathfrak{g}$.*

Proof. The assumption that the path $\alpha(t) \cdot x_0$ is deformable to a point implies the existence of a function $\xi(t, s)$ satisfying the conditions of Lemma 4 and such that $\xi(t, 1) = \alpha(t) \cdot x_0$. For each s, $\xi(t, s)$ is a path of the form $[x_0, x_0]$, and hence by Lemma 2 there exists a path defined, say, by $\alpha_s(t)$ beginning at e and such that $\alpha_s(t) \cdot x_0 = \xi(t, s)$. We may in particular take $\alpha_1(t) = \alpha(t)$. The family $\alpha_s(t)$ is not necessarily continuous in s. Now by Lemma 5 there exists for each $\bar{s}$ in the closed interval $[0, 1]$ an open interval $\sigma_{\bar{s}}$ such that, if s' is in $\sigma_{\bar{s}}$ and in $[0, 1]$, there will exist a path $\alpha_{s'}'(t)$ which can be deformed to $\alpha_{\bar{s}}(t)$, end points remaining fixed and such that $\alpha_{s'}'(t) \cdot x_0 = \xi(s', t)$. Let a finite covering subset of intervals, $\sigma_1, \cdots, \sigma_u, \sigma_v, \cdots, \sigma_0$, be chosen and arranged so that $1, \cdots, u, v, \cdots, 0$ is a decreasing sequence. There exists a w between u and v, and contained in both σ_u, σ_v. Hence there can be chosen paths $\alpha_w'(t)$ and $\alpha_w''(t)$, the first deformable to $\alpha_u(t)$ and the second to $\alpha_v(t)$, with end points always remaining fixed and such that $\alpha_w'(t) \cdot x_0 = \alpha_w''(t)$ $(= \xi(w, t))$. It follows by Lemma 3 that $\alpha_w'(t)$ can be deformed to $\alpha_w''(t)$ in such a way that the terminal point moves along a fixed coset of $\mathfrak{g}$, the initial point remaining fixed. Hence $\alpha_u(t)$ can be similarly deformed to $\alpha_v(t)$, and hence $\alpha_1(t)$ in a finite number of steps to $\alpha_0(t)$. Since the terminal point of $\alpha_1(t)$ is in $\mathfrak{g}$, it remains in $\mathfrak{g}$ during the deformation. Since $\alpha_0(t) \cdot x_0 = \xi(t, 0) = x_0$, the path $\alpha_0(t)$ is in $\mathfrak{g}$.

40. **Fundamental groups.** Let the fundamental group $G(\mathfrak{G})$ be defined by the paths beginning and ending at e, and $G(\mathfrak{X})$ by the paths beginning and ending at x_0. Under the correspondence $a \to a \cdot x_0$, the image of $G(\mathfrak{G})$ is a subgroup $G^0(\mathfrak{X})$ of $G(\mathfrak{X})$. Let $\mathfrak{g}_0$ be that connected part of $\mathfrak{g}$ which contains e, and let $G^0(\mathfrak{G})$ be the subgroup of $G(\mathfrak{G})$ consisting of all the paths which are equivalent to paths in $\mathfrak{g}_0$, that is, those paths that can be deformed to paths in $\mathfrak{g}_0$ through families of paths beginning and ending at x_0. We assert that

$$(1) \qquad G(\mathfrak{G})/G^0(\mathfrak{G}) = G(\mathfrak{X}).$$

Since the image of every path in $\mathfrak{g}_0$ is the point x_0, every element in $G^0(\mathfrak{G})$ corresponds to the identity in $G^0(\mathfrak{X})$. Conversely, every element of $G(\mathfrak{G})$

[23] This theorem is implied by Cartan in his remarks on fundamental groups ([2], p. 27), and again by Ehresmann ([3], p. 399).

which corresponds to the identity in $G^0(\mathfrak{X})$ is an element in $G^0(\mathfrak{G})$. To see this, let $\xi(t)$ $(0 \leqq t \leqq 1)$ be a path beginning and ending at x_0, and $\alpha(t)$ a path beginning and ending at e and such that $\alpha(t) \cdot x_0 = \xi(t)$. Suppose that $\xi(t)$ is deformable to x_0 through a family of paths beginning and ending at x_0. By the last theorem, $\alpha(t)$ is deformable to a path in $\mathfrak{g}$ through a family $\alpha_s(t)$ $[0 \leqq s \leqq 1, \alpha_0(t) = \alpha(t)]$ of paths beginning at e and ending in $\mathfrak{g}_0$. Let $\alpha'(s) = \alpha_s(1)$. Then $\alpha'(s)$ defines the path traced in $\mathfrak{g}_0$ by the terminal point of $\alpha(t)$. Let

$$\beta_s(t) = \begin{cases} \alpha_s\left(\dfrac{2t}{2-s}\right) & (0 \leqq t \leqq 1 - s/2), \\ \alpha'(-2t+2) & (1 - s/2 \leqq t \leqq 1). \end{cases}$$

It will be seen that $\beta_s(t)$ defines a deformation of $\alpha(t)$ to a path in $\mathfrak{g}_0$ through a family of paths beginning and ending at e. The resultant path in $\mathfrak{g}_0$ is a union of the paths defined by $\alpha_1(t)$ and $\alpha'(t)$. We have now established the relation (1).

41. Let the maximum number of linearly independent elements in an abelian group be called its *rank*. It is well known[24] that the fundamental group of a connected continuous group is abelian. In particular, the subgroup $G^0(\mathfrak{G})$ is abelian, and from its definition it is clear that its rank can not exceed the rank of $G(\mathfrak{g}_0)$. Moreover, the rank of the fundamental group of a connected continuous group of r dimensions[25] can not exceed r. Hence

$$\text{(1)} \qquad \operatorname{rank} G^0(\mathfrak{G}) \leqq \dim \mathfrak{g}_0 .$$

Moreover, it follows from (1), §40, that

$$\text{(2)} \qquad \operatorname{rank} G(\mathfrak{G}) = \operatorname{rank} G^0(\mathfrak{X}) + \operatorname{rank} G^0(\mathfrak{G}) ,$$

and from §35 that (since $\dim \mathfrak{g} = \dim \mathfrak{g}_0$)

$$\dim G = \dim \mathfrak{g}_0 + \dim \mathfrak{X} .$$

[24] Schreier, [10].

[25] Smith, [13]. Application of the results of [13] requires that $\mathfrak{G}$ be subdivisible into a simplicial complex. (Such a subdivision will automatically satisfy (*a*) and (*b*) of p. 210, since $\mathfrak{G}$ (as well as $\mathfrak{X}$) is a topological manifold (§36).) We can choose a coördinate system about e in which the functions $(ab)^i$ will be analytic (Schur [11]). Let σ be a spherical region with center at e, and let $a_1, a_2, \cdots$ be chosen such that $\Sigma\, a_i\sigma = \mathfrak{G}$ (§36). The cells $a_1\sigma, a_2\sigma, \cdots$ and their boundaries form a subdivision of $\mathfrak{G}$ into pieces of $r, r-1, \cdots$ dimensions, and if the radius of σ is small enough, this subdivision will be, at least locally, analytic in character. Hence it can be locally further subdivided into simplexes by the methods of [7] or [8]. There is no essential difficulty in coördinating the local simplicial subdivisions into a simplicial subdivision for the whole of $\mathfrak{G}$. The same remarks apply to $\mathfrak{g}_0$. With regard to $\mathfrak{X}$, we remark that among the coördinate systems for $\mathfrak{G}$ in which the $(ab)^i$ are analytic there are canonical systems. It then follows readily from §35 that, such a system being chosen for $\mathfrak{G}$, there exists a coördinate system in $\mathfrak{X}$ extending over a neighborhood of x_0 in which the $(a \cdot x)^i$ are analytic. By an argument very much like the one just outlined, this leads to a simplicial subdivision of $\mathfrak{X}$.

Hence we have

$$\text{rank } G(\mathfrak{G}) + \dim \mathfrak{X} \leqq \text{rank } G^0(\mathfrak{X}) + \dim \mathfrak{G} . \tag{3}$$

Thus, for example, *an r-parameter Lie group $\mathfrak{G}_r$ can not operate transitively in a euclidean space $\mathfrak{X}$ of n dimensions if the rank of the fundamental group of $\mathfrak{G}_r$ exceeds $r - n$.*

We assert finally that rank $G^0(\mathfrak{X}) \leqq \dim \mathfrak{X} = n$. For suppose, on the contrary, that there exist $n + 1$ paths $\alpha_1(t), \cdots, \alpha_{n+1}(t)$ beginning and ending at e such that the paths defined by $\alpha_1(t)\cdot x_0, \cdots, \alpha_{n+1}(t)\cdot x_0$ are linearly independent. The function

$$f(t_1, \cdots, t_n) = \alpha_1(t_1)\, \alpha_2(t_2) \cdots \alpha_n(t_n)\cdot x_0 , \tag{4}$$

where $t_1, \cdots, t_n$ are independent variables varying between 0 and 1, defines a single-valued continuous mapping of an n-dimensional torus τ on $\mathfrak{X}$ such that the n "parameter curves" of τ are mapped on the paths defined by $\alpha_1(t), \cdots, \alpha_n(t)$. Let $\tilde{\mathfrak{X}}$ be the covering space of $\mathfrak{X}$ which belongs[26] to the subgroup Γ of $G^0(\mathfrak{X})$ generated by $\alpha_1(t), \cdots, \alpha_{n+1}(t)$. The fundamental group of $\tilde{\mathfrak{X}}$ is[27] Γ, and is therefore abelian. Furthermore, an examination of the construction of $\tilde{\mathfrak{X}}$ shows easily that with the aid of (4) there can be defined a mapping of τ on $\tilde{\mathfrak{X}}$ such that the parameter curves are mapped on the first n generators of Γ. But since the rank of Γ is $n + 1$, such a mapping is impossible,[28] and our assertion follows. From (2) and the fact that rank $G^0(\mathfrak{G}) \leqq$ rank $G(\mathfrak{g}_0)$, we have rank $G(\mathfrak{G}) -$ rank $G(\mathfrak{g}_0) \leqq \dim \mathfrak{X}$. Thus, for example, if $\mathfrak{X}$ is n-dimensional and rank $G(\mathfrak{G}) = n + s$, then rank $G(\mathfrak{g}_0) \geqq s$.

Barnard College, Columbia University.

REFERENCES

[1] L. E. J. Brouwer, *On looping coefficients*, Amsterdam Proceedings, vol. 15 (1912), pp. 113–122.

[2] Elie Cartan, *La théorie des groupes finis et continus et l'analysis situs*, Mémorial des Sciences Mathématiques, no. 42.

[3] C. Ehresmann, *Sur la topologie de certains espaces homogènes*, Annals of Math., vol. 35(1934), pp. 396–443.

[4] L. P. Eisenhart, *Continuous Groups of Transformations*, Princeton University Press, 1933.

[5] F. Hausdorff, *Mengenlehre*, 1927.

[6] H. Hopf, *Zur Topologie der Abbildungen von Mannigfaltigkeiten*, II, Math. Ann., vol. 102 (1929).

[7] B. O. Koopman and A. B. Brown, *On the covering of analytic loci by complexes*, Transactions American Math. Society, vol. 34 (1932), pp. 231–251.

[8] S. Lefschetz and J. H. C. Whitehead, *On analytical complexes*, Transactions American Math. Society, vol. 35 (1933), pp. 510–517.

[26] See Hopf, [6], p. 568.

[27] Hopf, [6], p. 572.

[28] By the argument in [13], particularly p. 228.

[9] S. LIE AND F. ENGEL, *Theorie der Transformationsgruppen*, vol. 1 (1888) and vol. 3 (1893).
[10] O. SCHREIER, *Abstrakte continuierliche Gruppen*, Abh. math. Seminar, Hamburg, vol. 5 (1927), pp. 15–32.
[11] F. SCHUR, *Über den analytischen Character der eine endliche continuierliche Transformationsgruppe darstellenden Funktionen*, Math. Ann., vol. 41 (1893), pp. 519–538.
[12] P. A. SMITH, *Properties of group manifolds*, Proc. Nat. Acad., vol. 17 (1931), pp. 674–675.
[13] P. A. SMITH, *The fundamental group of a group manifold*, Annals of Math., vol. 36 (1935), pp. 210–229.

A CLASS OF QUATERNION ALGEBRAS

BY JAMES H. D. TELLER

1. **Introduction.** Let $\mathfrak{A}$ be a rational generalized quaternion algebra with basis elements $1, i, j, ij$, where $i^2 = -1, j^2 = \alpha, ij = -ji$. Without loss of generality we may take α odd and either ± 1, or a product of distinct primes of the form $4n + 3$ or the negative of such a product.[1] Latimer[2] has proved theorems similar to those of this paper, which include the case $\alpha \equiv 1 \pmod 4$. Hence we assume $\alpha \equiv 3 \pmod 4$. Then $\mathfrak{A}$ is a division algebra and the set $\mathfrak{G}$ of integral elements containing the basal elements $1, i, j, ij$ consists of all elements $x + \rho y$, where x, y range over the set G of Gaussian complex integers and $\rho = \eta(1 + j)$, where[3] $\eta = \frac{1}{2}(1 + i)$. The conjugate of $X = a + bi + cj + dij$ is $X' = 2a - X$ and the norm of X is $N(X) = XX' = a^2 + b^2 - \alpha c^2 - \alpha d^2$.

We shall show that there is a one-to-one correspondence between the classes of left ideals in $\mathfrak{G}$ and those classes of binary Hermitian forms $axx' + (b/2)\, x'y + (b'/2)xy' + cyy'$ which represent positive integers, where a, c are rational integers, b is in G, x and y range over G, b', x', y' are the conjugates of b, x, y respectively and $bb' - 4ac = 2\alpha$. If $\alpha > 0$, we prove two theorems on the existence of a g.c.d. and on the factorization of elements in $\mathfrak{G}$, almost identical with certain theorems proved by Dickson for the case $\alpha = -1$, i.e., the case where $\mathfrak{G}$ is the set of Hurwitz integral quaternions.

Latimer in the paper cited above proved similar theorems for the sets $x + jy$, where x, y range over the integral elements of a quadratic field. The corresponding forms in his case were $axx' + bx'y + b'xy' + cyy'$, where the same assumptions hold for the coefficients except that $bb' - ac = \alpha$. If b is a rational integer, and x, y are restricted to the set of rational integers, Latimer's forms become classic binary quadratic forms $ax^2 + 2bxy + cy^2$, while those of this paper become non-classic forms $ax^2 + bxy + cy$, b odd. Latimer's paper[4] extends the factorization theory of the Lipschitz integral quaternions; this paper makes a similar extension of the theory of the Hurwitz integral quaternions.

Received August 19, 1935.

[1] Dickson, *Algebras and Their Arithmetics*, p. 192.

[2] Latimer, *On ideals in generalized quaternion algebras and hermitian forms*, Transactions of the American Mathematical Society, vol. 38 (1935), pp. 436–446.

[3] Dickson, loc. cit., p. 192.

[4] Latimer, loc. cit.

2. **Basis of an ideal in $\mathfrak{G}$.** The following equations which may be verified are used in the sequel.

$$(1)\qquad \begin{aligned} &(a)\ \rho^2 = \rho - \epsilon,\\ &(b)\ N(\rho) = \epsilon,\\ &(c)\ x\rho = \rho x' + (x - x')\eta,\\ &(d)\ x\rho' = \rho' x' + (x - x')\eta',\\ &(e)\ N(\mathfrak{X}) = N(X) + \epsilon N(Y) + (x_1 + y_1)u_1 - (x_1 - y_1)v_1, \end{aligned}$$

where $\epsilon = (1 - \alpha)/2$, $\mathfrak{X} = X + \rho Y$, $X = x_1 + y_1 i$, $Y = u_1 + v_1 i$.

Ideal, equivalence of ideals, class of ideals, basis of an ideal with respect to G, and proper basis are defined as in Latimer's paper,[5] with E replaced by ρ, except that the phrase "non-singular" is replaced by "not equal to zero" as $\mathfrak{A}$ is a division algebra.

LEMMA 1. *Every ideal $\mathfrak{L}$ in $\mathfrak{G}$ has a basis a, $b + \rho d$ where a, b, d are in G.*

Let $a_i + \rho d_i$ be the set of elements of $\mathfrak{L}$. The set of d_i have a g.c.d. $d = \Sigma\, x_i d_i$, where the x's are in G. Then by (1c) $\Sigma\, x_i(a_i + \rho d_i) = b + \rho d$ is an element of $\mathfrak{L}$, where b is in G. Let the g.c.d. of the complex integers in $\mathfrak{L}$ be a. Then by employing (1), it will be found that a, $b + \rho d$ form a basis of $\mathfrak{L}$. We shall write $\mathfrak{L} = [a, b + \rho d]$.

LEMMA 2. *Every ideal $\mathfrak{L}$ in $\mathfrak{G}$ is equivalent to an ideal $[a, b + \rho]$, where a is an odd rational integer and b is in G.*

Let $\mathfrak{L}_1 = [a_1, b_1 + \rho d_1]$ be an ideal in $\mathfrak{G}$. Since ρa_1 and $(\rho - 1)(b_1 + \rho d_1) = -\epsilon d_1 - b_1 + \rho b_1$ are elements of $\mathfrak{L}_1$, $a_1 = a d_1$, $b_1 = b d_1$, where a, b are in G. Consider $\mathfrak{L} = [a, b + \rho]$, $\mathfrak{L} d_1 = \mathfrak{L}_1$. Since $N(d_1) > 0$, $\mathfrak{L}$ is equivalent to $\mathfrak{L}_1$.

It remains to show that a is an odd rational integer. Since $\rho a - a'(b + \rho) = (a - a')\eta - a'b$ and $(\rho - b' - 1)(b + \rho) = -\epsilon + (b - b')\eta - b(b' + 1)$ are in $\mathfrak{L}$, we have

$$(2)\qquad \begin{aligned} (a - a')\eta - a'b &\equiv 0 \pmod{a},\\ \epsilon - (b - b')\eta + b(b' + 1) &\equiv 0 \pmod{a}. \end{aligned}$$

Setting $b = b_1 + b_2 i$ in (2_2) we find

$$(3)\qquad (2b_1 + 1)^2 + (2b_2 + 1)^2 - 2\alpha = 4ac,$$

where c is in G. Since $\alpha \equiv 3 \pmod 4$, a and c are prime to 2. Let $a = t a_1$, where t is an odd rational integer and a_1 is in G and not divisible by a rational integer > 1. Then a_1' is prime to a_1. Multiplying the left member of (2_1) by $1 - i$ and noting that a_1' is prime to a_1 we find

$$(4)\qquad 1 + b(1 - i) \equiv 0 \pmod{a_1}.$$

Multiplying the left member of (2_2) by $1 - i$, reducing by means of (4) and then multiplying by $1 + i$, we find

$$(5)\qquad b(1 - i) + 2\epsilon \equiv 0 \pmod{a_1}.$$

[5] Latimer, loc. cit.

From (4) and (5),

$$1 - 2\epsilon = \alpha \equiv 0 \pmod{a_1}. \tag{6}$$

Since α has no non-rational complex prime factors, $a_1 = \pm 1, \pm i$. Hence we may assume $a_1 = 1$, and therefore a is an odd rational number.

If an ideal $\mathfrak{L}$ has a proper basis ω_1, ω_2; $\omega_i = g_{i1} + g_{i2}\rho$ $(i = 1, 2)$, where the determinant $|g_{ij}|$ is a positive rational integer, then $|g_{ij}|$ is defined to be the norm of $\mathfrak{L}$ and written $N(\mathfrak{L})$. It may be shown that $N(\mathfrak{L})$ is independent of the particular proper basis employed. If $\zeta_i = t_{i1}\omega_1 + t_{i2}\omega_2$ $(i = 1, 2)$, where the ω's form a proper basis of $\mathfrak{L}$, it may also be shown that the ζ's form a proper basis if and only if the t's are in G and $|t_{ij}| = 1$. By the proof of Lemma 2 every ideal $\mathfrak{L}$ in $\mathfrak{G}$ has a proper basis $[ad, bd + \rho d]$, where a is a positive odd rational integer and b, d are in G. Hence $N(\mathfrak{L}) = add'$. If ξ is an element of $\mathfrak{G}$ of positive norm, it may be shown that $N(\mathfrak{L}\xi) = N(\mathfrak{L})\,N(\xi)$.

3. **The class of forms corresponding to an ideal.** If a, c are rational integers, b is in G, x, y range over G and b', x', y' are the conjugates of b, x, y respectively, then

$$f(x, y) \equiv axx' + \tfrac{1}{2}bx'y + \tfrac{1}{2}b'xy' + cyy' \tag{7}$$

will be said to be a Hermitian form of discriminant $bb' - 4ac$. If $f_1(x_1, y_1)$ is obtained from f by a linear homogeneous transformation of determinant unity, f and f_1 will be said to be equivalent. Equivalent forms have equal discriminants. All the forms equivalent to a given form will be said to form a class.

Let $\mathfrak{X} = x + y\rho$ be an element of $\mathfrak{G}$. By (1) $\rho\mathfrak{X} = (x - x')\eta - y'\epsilon + [x' + y' + (y - y')\eta]\rho$. The determinant

$$\begin{vmatrix} x & y \\ (x - x')\eta - y'\epsilon & x' + y' + (y - y')\eta \end{vmatrix}$$

will be found equal to $N(\mathfrak{X})$ as given by (1e).

Let $\mathfrak{L}$ be an ideal with proper basis ω_1, ω_2; $\omega_i = g_{i1} + g_{i2}\rho$ $(i = 1, 2)$. Since each $\rho\omega_i$ belongs to $\mathfrak{L}$, we have

$$\rho\omega_i = b_{i1}\omega_1 + b_{i2}\omega_2 \qquad (i = 1, 2), \tag{8}$$

where the b's are in G. The general element of $\mathfrak{L}$ is $\mathfrak{X}$ as written below, where x, y range over G.

$$\mathfrak{X} = x\omega_1 + y\omega_2 = (g_{11}x + g_{21}y) + (g_{12}x + g_{22}y)\rho,$$

$$\rho\mathfrak{X} = l_1\omega_1 + l_2\omega_2 = (g_{11}l_1 + g_{21}l_2) + (g_{12}l_1 + g_{22}l_2)\rho,$$

where $l_1 = b_{11}x' + b_{21}y' + (x - x')\eta$, $l_2 = b_{12}x' + b_{22}y' + (y - y')\eta$. Then

$$N(\mathfrak{X}) = \begin{vmatrix} g_{11}x + g_{21}y & g_{12}x + g_{22}y \\ g_{11}l_1 + g_{21}l_2 & g_{12}l_1 + g_{22}l_2 \end{vmatrix} = \begin{vmatrix} x & y \\ l_1 & l_2 \end{vmatrix}\begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} = f(x, y)\,N(\mathfrak{L}),$$

where

$$(9)\qquad f(x, y) = \begin{vmatrix} x & y \\ l_1 & l_2 \end{vmatrix} = b_{12}xx' + (\eta - b_{11})x'y + (b_{22} - \eta)xy' - b_{21}yy'.$$

$f(x, y)$ will be said to correspond to the proper basis ω_1, ω_2 of $\mathfrak{L}$.

We have seen that $\zeta_i = t_{i1}\omega_1 + t_{i2}\omega_2$ $(i = 1, 2)$ form a proper basis if and only if the t's are in G and $|t_{ij}| = 1$. The form corresponding to such a basis is $f_1(x_1, y_1) = N(x_1\zeta_1 + y_1\zeta_2)/N(\mathfrak{L})$, where $\mathfrak{X} = x_1\zeta_1 + y_1\zeta_2 = x\omega_1 + y\omega_2$,

$$(10)\qquad \begin{aligned} x &= t_{11}x_1 + t_{21}y_1, \\ y &= t_{12}x_1 + t_{22}y_1. \end{aligned}$$

Hence f is equivalent to f_1. Conversely, if f is transformed into f_1 by a transformation (10), the t's being in G and $|t_{ij}| = 1$, then f_1 is the form corresponding to the proper basis $\zeta_1, \zeta_2, \zeta_i = t_{i1}\omega_1 + t_{i2}\omega_2$ $(i = 1, 2)$. Hence there is a one-to-one correspondence between the proper bases of $\mathfrak{L}$ and the forms in the class C containing f. We shall say that $\mathfrak{L}$ corresponds to C.

Multiplying $f(x, y)$ of (9) by 2, we have

$$2f(x, y) \equiv 2b_{12}xx' + [1 + i - 2b_{11}]\, x'y + [2b_{22} - 1 - i]\, xy' - 2b_{21}yy'.$$

Since $f(x, y) = N(\mathfrak{X})/N(\mathfrak{L})$ is rational, $2f$ is rational and is in G for every x, y in G, and therefore is a rational integer for every such x, y. It may then be shown that $2b_{12}, 2b_{21}$ are rational integers and that the coefficients of $x'y$ and xy' are conjugate complex integers. Since the b_{ij} are in G, b_{12} and b_{21} are rational integers and $f(x, y)$ of (9) is a form of type (7).

Lemma 3. *If C and C_1 are classes of hermitian forms which correspond to the ideals $\mathfrak{L}$ and $\mathfrak{L}_1$ respectively, then $C = C_1$ if and only if $\mathfrak{L}$ and $\mathfrak{L}_1$ are equivalent.*

Let $f(x, y)$ of (7) be a form in C. We may assume without loss of generality that $a \neq 0$. Suppose $C = C_1$. Then f corresponds to a proper basis ω_1, ω_2 of $\mathfrak{L}$ and to a proper basis ζ_1, ζ_2 of $\mathfrak{L}_1$. From (7), (8), (9) we have

$$\rho\omega_1 = b_{11}\omega_1 + b_{12}\omega_2 = (\eta - \tfrac{1}{2}b)\omega_1 + a\omega_2; \qquad \rho\zeta_1 = (\eta - \tfrac{1}{2}b)\zeta_1 + a\zeta_2$$

and

$$(\rho + \tfrac{1}{2}b - \eta)\, \omega_1 = a\omega_2, \qquad (\rho + \tfrac{1}{2}b - \eta)\zeta_1 = a\zeta_2.$$

From $N(x\omega_1 + y\omega_2) = N(\mathfrak{L})f(x, y)$ it follows that $N(\omega_1) = aN(\mathfrak{L})$. Similarly $N(\zeta_1) = aN(\mathfrak{L}_1)$. Then $N(\omega_1)\, N(\zeta_1) > 0$. We have

$$\mathfrak{L}a\omega_1' = [a\omega_1, a\omega_2]\, \omega_1' = [a, \rho + \tfrac{1}{2}b - \eta]\, N(\omega_1).$$

Similarly, $\mathfrak{L}_1a\zeta_1' = [a, \rho + \tfrac{1}{2}b - \eta]\, N(\zeta_1)$. Then $\mathfrak{L}a\omega_1'N(\zeta_1) = \mathfrak{L}_1a\zeta_1'N(\omega_1)$, and $\mathfrak{L}$ is equivalent to $\mathfrak{L}_1$.

Conversely, suppose $\mathfrak{L}\xi = \mathfrak{L}_1\xi_1$, and $N(\xi)N(\xi_1) > 0$. We have $\mathfrak{L}\xi = [\omega_1\xi, \omega_2\xi] =$

$[\zeta_1\xi_1, \zeta_2\xi_1]$ where the ω's and ζ's form proper bases of $\mathfrak{L}$ and $\mathfrak{L}_1$, respectively. Since we may assume that ξ is a rational integer and $N(\xi_1) > 0$,

$$\omega_1\xi = \xi(g_{11} + g_{12}\rho),$$

$$\omega_2\xi = \xi(g_{21} + g_{22}\rho),$$

and the first basis of $\mathfrak{L}\xi$ is proper. Moreover, if $\zeta_i = h_{i1} + h_{i2}\rho$ $(i = 1, 2)$, by (1c)

$$\zeta_1\xi_1 = (h_{11} + h_{12}\rho)\xi_1, \qquad \zeta_2\xi_1 = (h_{21} + h_{22}\rho)\xi_1,$$

and it will be found that these elements form a proper basis of $\mathfrak{L}\xi$. Hence $\omega_i\xi = (t_{i1}\zeta_1 + t_{i2}\zeta_2)\xi_1$ $(i = 1, 2)$, where the t's are in G and $|t_{ij}| = 1$. Let $f(x, y)$ be the form in C corresponding to the proper basis ω_1, ω_2 of $\mathfrak{L}$, and let $f_1(x_1, y_1)$ be the form in C_1 corresponding to the proper basis ζ_1, ζ_2 of $\mathfrak{L}_1$. Then $N(\mathfrak{L})N(\xi)f(x, y) = N[(t_{11}x + t_{21}y)\zeta_1\xi_1 + (t_{12}x + t_{22}y)\zeta_2\xi_1] = N(\mathfrak{L}_1)f_1(x_1, y_1)N(\xi_1)$. But $N(\mathfrak{L})N(\xi) = N(\mathfrak{L}\xi) = N(\mathfrak{L}_1\xi_1) = N(\mathfrak{L}_1)N(\xi_1) \neq 0$. Therefore $f(x, y) = f_1(x_1, y_1)$ and $C = C_1$.

4. **The correspondence between classes of ideals and classes of forms.** We shall prove

THEOREM 1. *There is a one-to-one correspondence between the classes of ideals in $\mathfrak{G}$ and those classes of hermitian forms* (7) *of discriminant 2α which represent positive integers.*

By Lemma 3, for every class of ideals there is a uniquely determined class of forms. Also, no class of forms corresponds to two classes of ideals. To prove the above theorem, it is therefore sufficient to show that (a) if C is a class of forms corresponding to a class of ideals, then C contains a form which represents a positive integer and is of discriminant 2α, (b) every class of forms of discriminant 2α which represents a positive integer corresponds to a class of ideals in $\mathfrak{G}$.

From Lemma 2 every class of ideals contains an ideal $\mathfrak{L} = [a, b + \rho]$, where a is a positive odd rational integer and b is in G. The indicated basis of $\mathfrak{L}$ is proper, $N(\mathfrak{L}) = a$. From (8) and (2) we get $b_{11} = -b$, $b_{12} = a$, $b_{21} = [-\epsilon + (b - b')\eta - b(b' + 1)]/a = -c$, $b_{22} = b' + 1$. Substituting these values in (9) we obtain $f(x, y) = axx' + [(2b + 2\eta)/2]x'y + [(2b' + 2\eta')/2]xy' + cyy'$. Then f represents the positive integer a, the discriminant of f is 2α and the class containing f corresponds to the class containing $\mathfrak{L}$. This proves (a). Let C be a class of forms of discriminant 2α which represent a positive integer a_1 and let f of (7) be a form in C. Then for properly chosen x, y, which we may assume are relatively prime, $f(x_1, y_1) = a_1$. Then f is equivalent to a form with leading coefficient a_1. Hence we may assume $a > 0$. Since $bb' = 4ac + 2\alpha \equiv 2 \pmod 4$, $b_1 = (b - 2\eta)/2$ is in G. Then it may be shown that there is an ideal $\mathfrak{L} = [a, b_1 + \rho]$ which corresponds to f. If $\mathfrak{X} = ax + y(b_1 + \rho)$ is the general element of $\mathfrak{L}$, $N(\mathfrak{X}) = af(x, y)$. Since $a > 0$, the above basis of $\mathfrak{L}$ is proper and C is the class corresponding to the class of ideals containing $\mathfrak{L}$.

5. **A class of algebras in which every ideal is principal.** A principal ideal $\{\eta\}$ is the set of all elements $\xi\eta$, where ξ ranges over $\mathfrak{G}$ and η is a fixed element in $\mathfrak{G}$.

We shall show that if $\alpha > 0$, every ideal in $\mathfrak{G}$ is principal. Since α contains no square factors, f of (7) is a primitive form. If in addition $\alpha > 0$, f is an indefinite form, and it may be shown that a, c, b_1, and b_2 are all odd. We shall prove

THEOREM 2. *Every indefinite primitive form* $f \equiv axx' + \frac{1}{2}bx'y + \frac{1}{2}b'xy' + cyy'$ *represents* 1.

We may assume $a > 0$, $c < 0$, a prime to c. For suppose $0 < a \leqq c$, $(a, c) = d$. Then the transformation $x = x_1 + ky_1$, $y = y_1$ carries f into $f_1 \equiv ax_1x_1' + \frac{1}{2}B\,x_1'y_1 + \frac{1}{2}B'x_1y_1' + Cy_1y_1'$, where $B = 2ak + b$, and if we set $k = k_1 + k_2i$, $b = b_1 + b_2i$, then $C = a(k_1^2 + k_2^2) + b_1k_1 + b_2k_2 + c$. By proper choice of k_1, k_2, the number $b_1k_1 + b_2k_2 + c$ may be made prime to a, while $|2ak_1 + b_1| \leqq a$, $|2ak_2 + b_2| \leqq a$. Then C is prime to a. If we set $B = B_1 + B_2i$ and note that $|B_i| = |2ak_i + b_i|$, then the discriminant of f is $BB' - 4ac = B_1^2 + B_2^2 - 4aC = 2\alpha > 0$. But $B_1^2 + B_2^2 \leqq 2a^2 < 4a^2$, so that $B_1^2 + B_2^2 - 4a^2 < 0$. Hence $C < a$. If $C > 0$ we may set $x = -y_1$, $y = x_1$, and then repeat this process. We eventually get a form f of type (7) with $a > 0$, $c < 0$ and a prime to c.

The transformation $x = 2\,(nz + it)$, $y = u + imt$ carries f into

$$\phi(z, t, u) \equiv Hz^2 + 4at^2 + cu^2 + 2b_1nzu + 2b_1mzt + 2b_2tu,$$

a ternary form of determinant $D = -2\alpha H$, where $H \equiv 4an^2 + cm^2 - 2b_2mn$. H is an indefinite primitive binary quadratic form and for proper choice of m, n represents k, the negative of a prime, prime to 2α. Then ϕ is a primitive indefinite form. The adjoint Φ of ϕ is a ternary form of determinant D^2. Since D contains no square factors, Φ is a primitive form. The coefficient of z^2 in Φ is $4ac - b_2^2$, an odd negative number. Hence ϕ and Φ are both properly primitive indefinite forms. Then Φ is the negative of the reciprocal of ϕ, and the invariants are[6] $\Omega = -1, \Delta = D$. Since $D \equiv 2$ or $6 \pmod 8$, a binary quadratic of determinant D has $h + 1$ characters where h is the number of odd prime divisors of D. Hence f represents 1.[7]

Therefore every form (7) is equivalent to a form with leading coefficient unity. To every such form by the second part of Theorem 1 corresponds the principal ideal $[1, b_1 + \rho] = \{1\}$. Hence every class of ideals is equivalent to $\{1\}$. Since it may be shown that every ideal equivalent to a principal ideal is itself principal, we conclude that every ideal in $\mathfrak{G}$ is principal.

6. **Existence of a g.c.r.d. and factorization of elements of $\mathfrak{G}$.** Employ the definitions of unit and greatest common right divisor given by Latimer.[8]

[6] Dickson, *Studies in the Theory of Numbers*, p. 10.
[7] Dickson, loc. cit., p. 63, Theorems 52, 54 (with $m = 1$).
[8] Latimer, loc. cit.

We have proved the first sentence in

THEOREM 3. *If $\alpha > 0$, every ideal in $\mathfrak{G}$ is principal. Let λ, μ be elements in $\mathfrak{G}$, $\lambda \neq 0$; then λ, μ have a g.c.r.d. δ, which is uniquely determined apart from a unit left factor, and $\delta = \xi\lambda + \eta\mu$ where ξ, η are in $\mathfrak{G}$. If λ has no rational prime factor and $N(\lambda) = \pm p_1 p_2 \cdots p_r$, where the p's are rational primes arranged in an arbitrary but fixed order, then $\lambda = \pi_1\pi_2 \cdots \pi_r$, where $N(\pi_i) = \pm p_i$ and each π_i is uniquely determined apart from a unit left factor.*

The proof of the remaining part of this theorem may be made word for word as in Latimer's proof[9] of his Theorem 4 starting from the point "$\cdots$ $\mathfrak{L}$ is a principal ideal $\{\lambda, \mu\} = \{\delta\} \cdots$", except that the phrase "non-singular" is to be replaced by "not equal to zero".

Since every ideal is equivalent to $\{1\} = [1, \rho]$ for the case where $\alpha > 0$, it follows that every form f of (7) is equivalent to $N(x + \rho y) = F(x, y)$. Since F obviously represents 2, to show that every such form is universal it is sufficient to show that F represents -1 and every odd prime. It may be shown that F represents -1 by an argument similar to that used to show that it represents 1. It is well known that if p is an odd prime, there are rational integers a, b such that $a^2 + b^2 - \alpha \equiv 0 \pmod{p}$. If we set $\mu = a + bi + j$, it may then be shown as in the proof of Theorem 3 that the set of all elements $\xi p + \eta\mu$, where ξ, η range over $\mathfrak{G}$, form an ideal $\{\delta\}$, where $N(\delta) = \pm p$. Hence every such form f is universal.

We have seen that there is a single class of left ideals in $\mathfrak{G}$, if $\alpha > 0$. By a result due to Brandt[10] it follows that there is a single class of left ideals for every set $\mathfrak{G}_1$ of integral elements in $\mathfrak{A}$. We may then deduce the same results on the existence of a g.c.d. and factorization for any set $\mathfrak{G}_1$ in $\mathfrak{A}$ of integral elements as for the $\mathfrak{G}$ treated here.

UNIVERSITY OF KENTUCKY.

[9] Latimer, loc. cit.

[10] Brandt, *Idealtheorie in Quaternionenalgebren*, Mathematische Annalen, vol. 99 (1928), p. 23.

CONVERGENCE OF SEQUENCES OF POSITIVE LINEAR FUNCTIONAL OPERATIONS

By R. P. Bailey

Introduction. L. Fejér has recently called attention to the importance, in certain convergence problems of analysis, of a class of functionals to which he gives the name *positive operations.*[1] By his definition (loc. cit., p. 523), *a functional $U(x)$ defined over a set of functions $\{x(t)\}$, real-valued throughout a certain fundamental range $a \leqq t \leqq b$, is said to be positive, provided $U(x) \geqq 0$ whenever $x(t) \geqq 0$ in (a, b).* Sequences of operations of this kind often arise in the singular integral theory, in interpolation and in the theory of mechanical quadratures; in certain particular cases their convergence properties have been the object of much investigation. In his classical paper of 1909, Lebesgue gave the sequences of positive functionals which occur in the singular integral theory a special treatment, emphasizing again and again the simplicity of the reasoning involved, and their comparatively wide convergence properties.[2] At various times, many other writers have pointed out simplifications in a general theory which result from the hypothesis that certain sequences of functionals involved have the positive property.

It is our purpose, in the first part of this paper, to apply to the special case of the positive operations the very general theory which Hahn,[3] Banach[4] and others have developed for convergence problems involving sequences of linear functionals, with the object in view of establishing a set of theorems from which the particular theorems of Fejér, Lebesgue and others in the singular integral theory, mechanical quadratures and interpolation will be a matter of direct inference. The main discussion is divided into three parts. We take up first the question of the convergence of a sequence of positive linear functionals $\{U_n(x)\}$ $(n = 1, 2, \cdots)$ to the value of the function $x(t)(n \to \infty)$ at a certain fixed point $t = \tau$ of the fundamental interval. Sequences of this kind are familiar in interpolation and in the singular integral theory. In the second

Received September 26, 1935. The author wishes to express his gratitude to Professor J. A. Shohat for many valuable suggestions which have aided in the preparation of this paper.

[1] L. Fejér, *On the infinite sequences arising in the theories of harmonic analysis, interpolation and mechanical quadratures*, Bulletin of the American Mathematical Society, vol. 39 (1933), pp. 521–534.

[2] H. Lebesgue, *Sur les intégrales singulières*, Annales de Toulouse, (3), vol. 1(1909), pp. 25–117.

[3] H. Hahn, *Über Folgen linearer Operationen*, Monatshefte für Mathematik und Physik, vol. 32 (1922), pp. 3–88.

[4] S. Banach, *Théorie des Opérations Linéaires*, Monografje Matematyczne, Warsaw, 1932, pp. 122–130.

part we consider the convergence of the above sequence to a limit functional having the absolute continuity property of the Riemann integral and make applications to the problem of mechanical quadratures. In the third part we discuss an extension of the property of positiveness to more general operations, together with certain results on uniform convergence which can be obtained in this way.

The last section is devoted to a proof of the fact that one of the most important, although very special, types of positive operation, namely, the formulas of mechanical quadratures of Gauss' type, can have equal Côtes coefficients only in the well-known case of the trigonometric polynomials.

1. **Definition and properties of positive linear functional operations. Sequences of positive functionals.** Consider the function-space G whose elements $\{x(t)\}$ are the set of real-valued bounded functions of a real variable t defined over a closed finite interval[5] $a \leqq t \leqq b$. Two points of G are considered distinct if the corresponding functions differ at least at one point of (a, b). With the usual definitions of sum and scalar product, this set constitutes a linear or vectorial space. The zero element $x = \Theta$ is the point corresponding to the identically vanishing function $x(t) \equiv 0$. The point corresponding to the function $x(t) \equiv 1$ we denote by $x = I$. The distance (x, y) between two arbitrary points x and y of G we define to be $\underset{a \leqq t \leqq b}{\text{L.U.B.}} \mid x(t) - y(t) \mid$. With this distance-function, G becomes a metric space. It may be considered normal if we define the norm $\| x \|$ of a point x of G to be the non-negative number (x, Θ). If x and y belong to G and $x(t) \geqq y(t)$, we write $x \geqq y$.

A sequence of points $\{x_n\} (n = 1, 2, \cdots)$ of G is said to converge to the point x_0 of G, provided $\lim_{n \to \infty} \| x_n - x_0 \| = 0$. This clearly requires the uniform convergence of the sequence of functions $\{x_n(t)\}$ to $x_0(t)$.

A functional operation $U(x)$ defined over a vectorial subset E of G associates with each x of E a real number $U(x)$. This operation is said to be linear, if

$$\text{(i)} \quad U(x_1 + x_2) = U(x_1) + U(x_2) \qquad (x_1, x_2 \subset E),$$

$$\text{(ii)} \quad \lim_{n \to \infty} U(x_n) = U(x_0), \text{ whenever } \| x_n - x_0 \| \to 0 \qquad (\{x_n\}, x_0 \subset E);$$

that is, a linear operation is *additive* and *continuous*. Every linear operation is *homogeneous*, i.e., for an arbitrary real constant c and every x of E, $U(cx) = cU(x)$. Every linear operation is necessarily *bounded*, i.e., there exists a positive number $\mid U \mid_E = \underset{x \subset E, \|x\| \leqq 1}{\text{L.U.B.}} \mid U(x) \mid$, called the *norm* of the operation U over E, such that $\mid U(x) \mid \leqq \mid U \mid_E \| x \|$ for every x of E. As was stated above, U is said to be *positive* over E, if $U(x) \geqq 0$ whenever $x \geqq \Theta$. It follows at once, in virtue of (i), that here $x_1 \geqq x_2$ implies $U(x_1) \geqq U(x_2)$.

[5] This interval will be understood to be fixed throughout the paper unless the contrary is explicitly stated. In terminology, definitions and notations we follow S. Banach (loc. cit.).

The norm of a positive linear functional frequently has a very simple expression, as we show immediately.

LEMMA 1. *If the linear functional operation $U(x)$ is positive over a vectorial subset E of G which contains the point $x = I$, then $| U |_E = U(I)$.*

Evidently if $\| x \| \leqq 1$, $-I \leqq x \leqq I$, and therefore, since $U(x)$ is positive over E, $-U(I) \leqq U(x) \leqq U(I)$ and $| U(x) | \leqq U(I)$; whence

$$| U |_E = \underset{x \subset E, \| x \| \leqq 1}{\text{L.U.B.}} | U(x) | = U(I).$$

We shall need this result later on.

We now turn to the convergence of sequences of functionals. In general, for the convergence of a sequence $\{U_n(x)\} (n = 1, 2, \cdots)$ of linear functionals over a space E, it is sufficient to know (i) that it converges over some subspace H of E which is dense in E, and (ii) that the set of norms $\{| U_n |_E\}$ is bounded.[6] However, in the case of positive operations, these two hypotheses are sometimes redundant. Under certain conditions, (i) implies (ii), and hence (i) alone is sufficient. This point is brought out by

THEOREM 1. *Let E be a vectorial subset of G which contains an element of positive lower bound in (a, b). If a sequence $\{U_n(x)\} (n = 1, 2, \cdots)$ of positive linear functionals defined over E converges over a subset H dense in E to a linear[7] functional $U(x)$ defined over E, then $\lim_{n \to \infty} U_n(x) = U(x)$ for every x of E.*

It suffices to show that the set of norms $\{ | U_n |_E\} (n = 1, 2, \cdots)$ is bounded. Since E contains an element of positive lower bound in (a, b), and H is dense in E, then H necessarily contains a positive element bounded away from zero, say $\xi(t) \geqq \rho > 0$. The function $\eta(t) = \rho^{-1}\xi(t)$ now has the property that $\eta(t) \geqq 1$, whence $-\eta \leqq x \leqq \eta$ for all x of E such that $\| x \| \leqq 1$. It follows, since the operations involved are positive, that $-U_n(\eta) \leqq U_n(x) \leqq U_n(\eta)$ and $| U_n(x) | \leqq U_n(\eta)$ $(\| x \| \leqq 1;\ n = 1, 2, \cdots)$, whence $| U_n |_E \leqq U_n(\eta)$ $(n = 1, 2, \cdots)$. This shows that the set of norms $\{| U_n |_E\}$ is bounded, since the sequence converges for $x = \xi$, and therefore for $x = \eta$.

That some restriction of the kind we have made on the subset E is necessary for the validity of the theorem can be seen from the following example. The functional

$$\sigma_n(x) = \frac{1}{2n\pi} \int_0^{2\pi} x(t) \left\{ \frac{\sin \dfrac{nt}{2}}{\sin \dfrac{t}{2}} \right\}^2 dt$$

represents the arithmetic mean of the first n partial sums, at the point $t = 0$, of the Fourier series associated with the integrable function $x(t)$, of period 2π. In particular,

[6] Banach, loc. cit., p. 123.

[7] Under very general conditions, the limit functional of a convergent sequence of linear functionals is necessarily linear itself. Cf. Banach, loc. cit., p. 122.

$$\sigma_n(I) = \frac{1}{2n\pi}\int_0^{2\pi}\left\{\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}}\right\}^2 dt = 1 \qquad (n = 1, 2, \cdots). \tag{1}$$

Suppose now that our fundamental interval (a, b) is $(0, 2\pi)$. Let E be the set of continuous periodic functions $\{x(t)\}$ such that $x(0) = 0$, and let H denote that subset of E which consists of the trigonometric sums $\sum_{i=0}^{k} a_i \cos it + b_i \sin it$ $(k = 0, 1, 2, \cdots)$ which are zero at $t = 0$. H is dense in E by the Weierstrass theorem. Nevertheless, it can be shown that *the sequence of positive linear functionals*

$$U_n(x) \equiv n^{\frac{1}{2}}\sigma_n(x) \qquad (n = 1, 2, \cdots) \tag{2}$$

converges to the linear functional $U(x) = x(0) = 0$ *over* H *without converging at all points of* E. In fact, if x belongs to H, since the ordinary partial sums $S_n(x)$ of the first $2n + 1$ terms of the Fourier series associated with a trigonometric sum of order k are identical with the trigonometric sum itself, for $n \geqq k$, we readily conclude that in this case[8]

$$\lim_{n\to\infty} n^{1-\epsilon}\sigma_n(x) = 0$$

for any $\epsilon > 0$, and in particular for $\epsilon = \frac{1}{2}$. This shows that the sequence (2) converges over H. However, by (1),

$$| U_n |_E = \frac{n^{\frac{1}{2}}}{2\pi n}\int_0^{2\pi}\left\{\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}}\right\}^2 dt = n^{\frac{1}{2}} \qquad (n = 1, 2, \cdots),$$

and therefore the sequence of norms $\{| U_n |_E\}$ is *unbounded*. Since E is complete in this case, we may conclude that there exists a point x of $E - H$ at which the sequence cannot converge to the value $x(0) = 0$, for it is well known that a sequence of linear functionals cannot converge over a complete vectorial space unless the norms are bounded in their set.[9] Theorem 1 does not apply here, since E does not contain an element bounded away from zero in $(0, 2\pi)$.

In what follows we shall find it convenient to refer to certain particular subsets of G as follows: P = the set of all polynomials of degrees $0, 1, 2, \cdots$, with real coefficients; T = the set of all trigonometric sums $\sum_{i=0}^{k} a_i \cos it + b_i \sin it$ $(k = 0, 1, 2, \cdots)$ with real coefficients; C = the set of all continuous functions; R = the set of all bounded R-integrable functions; L = the set of all bounded L-integrable functions; S = the set of step-functions with a finite

[8] $n^{1-\epsilon}\sigma_n = n^{-\epsilon}[S_0 + S_1 + \cdots + S_{n-1}] = n^{-\epsilon}[S_0 + S_1 + \cdots + S_{k-1}] = o(1)$, since $S_k = S_{k+1} = \cdots = S_{n-1} = 0$.

[9] Banach, loc. cit., p. 80, Th. 5.

number ($\geqq 0$) of steps; K = the set of all functions of bounded variation. It should be remarked that the same metric (that of G) is used throughout.

If E is an arbitrary subset of G, the symbols E_τ and E_p will be used to denote respectively the set of functions $\{x(t)\}$ of E which are continuous at a certain fixed point $t = \tau$ of (a, b), and those functions of E which are periodic, of period 2π. The symbol $E_{\tau p}$ will denote the set of functions of E having both properties.

With these notations, the subsets H and E of Theorem 1 may be taken to be P and C respectively, for C is evidently a vectorial subset of G containing positive elements bounded away from zero, in which the polynomials are dense by the Weierstrass theorem. Similarly, we might take $H = T$ and $E = C_p$, or $H = S$ and $E = K$.

2. **The limit-functional** $U(x) = x(\tau)$. When the convergence of a sequence of positive linear functionals $\{U_n(x)\}$ over a subset H dense in a vectorial subset E of G is known, Theorem 1 will, in general, enable us to draw conclusions about the convergence of the sequence at other points of E. However, in the particular case where the limit-functional is $U(x) = x(\tau)$, τ being some fixed point of (a, b), conclusions can be drawn about the convergence of the sequence over certain subsets of G from much weaker hypotheses. Theorem 2, below, illustrates this point. We first establish

LEMMA 2. *Let u, l denote, respectively, the upper and lower bounds of the bounded function $x(t)$ at the point $t = \tau$. Then there exist continuous functions $\xi_1(t)$ and $\xi_2(t)$ such that* (i) $\xi_1 \leqq x \leqq \xi_2$, (ii) $\xi_1(\tau) = l$, $\xi_2(\tau) = u$.

Let us construct, for example, the function ξ_2. We can suppose that $u < \| x \|$; if $u = \| x \|$, evidently $\xi_2(t) \equiv u$ is the function required. Let $\delta > 0$ be so chosen that $\underset{\tau-\delta \leqq t \leqq \tau+\delta}{\text{L.U.B.}} x(t) < \| x \|$, and denote by u_k $(k = 0, 1, 2, \cdots)$ the least upper bound of $x(t)$ in the interval $(\tau - \delta/2^k, \tau + \delta/2^k)$. Clearly $u_0 \geqq u_1 \geqq u_2 \geqq \cdots \geqq u_k \geqq \cdots$; $\lim_{k\to\infty} u_k = u$. Denote by P_k, Q_k respectively the points whose coördinates are $(\tau - \delta/2^{k+1}, u_k)$, $(\tau + \delta/2^{k+1}, u_k)$ and by P the point whose coördinates are (τ, u). Evidently $\lim_{k\to\infty} P_k = \lim_{k\to\infty} Q_k = P$; hence the polygonal line joining the points $(a, \| x \|)$, $(\tau - \delta, \| x \|)$, $P_0, P_1, P_2, \cdots$ in succession from the left, and the points $(b, \| x \|)$, $(\tau + \delta, \| x \|)$, $Q_0, Q_1, Q_2, \cdots$ in succession from the right, is continuous throughout (a, b). It defines a continuous function $\xi_2(t)$ which evidently meets the requirements of Lemma 2. The function $\xi_1(t)$ can be constructed in a similar manner. It should be noted that *if $x(t)$ is continuous at $t = \tau$, then $\xi_1(\tau) = \xi_2(\tau) = x(\tau)$.*

THEOREM 2. *Let E be a vectorial subset of G containing $C(C_p)$. If a sequence $\{U_n(x)\}(n = 1, 2, \cdots)$ of positive linear functionals defined over E converges over a subset H dense in $C(C_p)$ to the value $x(\tau)$ [τ a fixed point of (a, b)], then $\lim_{n\to\infty} U_n(x) = x(\tau)$ for every x of $E_\tau(E_{\tau p})$.*

Consider the first case; let x be any point of E. Since $x(t)$ is continuous at

$t = \tau$, we may conclude (by Lemma 2) that there exist functions $\xi_1(t)$, $\xi_2(t)$, belonging to C, such that

$$\xi_1 \leqq x \leqq \xi_2, \tag{3}$$

$$\xi_1(\tau) = x(\tau) = \xi_2(\tau). \tag{4}$$

Since the operations $\{U_n(x)\}$ are positive over E, (3) implies that

$$U_n(\xi_1) \leqq U_n(x) \leqq U_n(\xi_2) \qquad (n = 1, 2, \cdots). \tag{5}$$

By Theorem 1 both $\lim_{n\to\infty} U_n(\xi_1) = \xi_1(\tau)$ and $\lim_{n\to\infty} U_n(\xi_2) = \xi_2(\tau)$ exist, and hence by (5) $\xi_1(\tau) \leqq \overline{\underline{\lim}}_{n\to\infty} U_n(x) \leqq \xi_2(\tau)$. This proves our statement, in virtue of (4). When C and E_τ are replaced by C_p and $E_{\tau p}$, respectively, ξ_1 and ξ_2 must be so chosen as to be periodic, of period 2π, as well as continuous. Though Lemma 2 does not affirm the existence of such functions, it is clear that methods analogous to those used in its proof will furnish the required construction, if we use the fact that $x(t)$ (belonging to $E_{\tau p}$) is now periodic, of period 2π, as well as continuous at $t = \tau$.

The question naturally arises whether a sequence of positive linear functionals $\{U_n(x)\}$ satisfying the hypotheses of Theorem 2 will not further converge to the value $\frac{1}{2}[x(\tau + 0) + x(\tau - 0)]$ for those functions $\{x(t)\}$ of E for which this functional is defined. This cannot be the case, in general, as one sees easily by considering the special sequence

$$U_n(x) \equiv x(\tau + 0) \qquad (n = 1, 2, \cdots). \tag{6}$$

The sequence (6) is defined, for instance, over K, and converges to the value $x(\tau)$ over K_τ by identity, but cannot converge to the value $\frac{1}{2}[x(\tau + 0) + x(\tau - 0)]$ at any point of K where $x(\tau + 0) \neq x(\tau - 0)$.

The following theorem gives a sufficient condition for convergence of the type described above.

THEOREM 3. *Let E be a vectorial subset of G containing K. If a sequence $\{U_n(x)\}(n = 1, 2, \cdots)$ of positive linear functionals defined over E converges over S to the value $\frac{1}{2}[x(\tau + 0) + x(\tau - 0)]$, then $\lim_{n\to\infty} U_n(x) = \frac{1}{2}[x(\tau + 0) + x(\tau - 0)]$ for every x of E which has a discontinuity of the first kind at $t = \tau$.*

The proof follows the same lines as that of Theorem 2. Since S is dense in K, Theorem 1 assures the convergence of the sequence to the required value over K. We now have only to construct, by the methods of Lemma 2, two functions ξ_1, ξ_2 of K such that $\xi_1 \leqq x \leqq \xi_2$ and $\xi_1(\tau + 0) = \xi_2(\tau + 0) = x(\tau + 0)$, $\xi_1(\tau - 0) = \xi_2(\tau - 0) = x(\tau - 0)$. The convergence of the sequence for the element x (of which we assume only that it has at most a discontinuity of the first kind at $t = \tau$) follows as in Theorem 2.

We may mention in passing an obvious extension of the theorems of this section which can be proved by means of Lemma 2.

THEOREM 4. *Let u_x, l_x denote respectively the upper and lower bounds of $x(t)$*

at the point $t = \tau$. *If a sequence* $\{U_n(x)\}(n = 1, 2, \cdots)$ *of positive linear functionals defined over a vectorial subset* E *of* G *which contains* C *converges over a subset* H *dense in* C *to the value* $x(\tau)$, *then* $l_x \leqq \underline{\lim}_{n\to\infty} U_n(x) \leqq \overline{\lim}_{n\to\infty} U_n(x) \leqq u_x$ *at every point of* E.

In Theorem 2 we may set $E = L$, $H = S_\tau$ (the set of step-functions with a finite number of steps which are continuous at $t = \tau$), and

$$U_n(x) \equiv \int_a^b x(t)\varphi(t - \tau, n)\, dt,$$

where, for all $n = 1, 2, \cdots, \varphi(\alpha, n)$ is a bounded, non-negative, L-integrable function of α. These substitutions are admissible, and our conclusion in this special case is that the convergence of the given sequence (a so-called "singular" integral) to the value $x(\tau)$ over S_τ implies convergence to the value $x(\tau)$ for every bounded L-integrable function $x(t)$ continuous at $t = \tau$. This result was obtained for the first time by Lebesgue.[10]

The choice of the set H of Theorem 2 depends, in general, upon the nature of the functionals $\{U_n(x)\}$. For instance, in discussing the convergence properties of Fejér's integral, we might take $H = T$ rather than $H = S_\tau$ (as in the theorem of Lebesgue), for the convergence of the Fejér sequence over T follows directly from its definition, while the question of its convergence over S_τ requires further investigation.

In a similar manner, Theorem 3 and Theorem 4 can be interpreted so as to give certain results of Lebesgue and others on the behavior of singular integrals for functions discontinuous[11] at $t = \tau$.

The following is an application of Theorem 2 in a situation where the singular integral theory cannot be used directly, i.e., interpolation.[12] Let

$$\begin{array}{lllll}
t_{11} & & & & \\
t_{12} & t_{22} & & & \\
t_{13} & t_{23} & t_{33} & & \\
\cdots & \cdots & \cdots & & \\
t_{1n} & t_{2n} & t_{3n} & \cdots & t_{nn} \\
\cdots & \cdots & \cdots & \cdots & \cdots
\end{array}$$

[10] Lebesgue, loc. cit., p. 75. The hypothesis of Lebesgue requires the convergence of the given sequence of integrals over what is known as the base of Hamel of the set S_τ (i.e., a linearly independent subset of S_τ upon which every other element of S_τ is dependent) and therefore is equivalent to an assumption of convergence over S_τ itself. In the same way, instead of assuming that a certain sequence of linear functionals converges for all polynomials, we may assume convergence only over the set $1, t, t^2, \cdots$; the two assumptions are equivalent, since each is an immediate inference of the other.

[11] Lebesgue, loc. cit., p. 78. See also Hobson, *Theory of the Functions of a Real Variable*, vol. 2(1926), p. 456, where a special case of Theorem 4 is discussed.

[12] H. Hahn, *Über Interpolation*, Mathematische Zeitschrift, vol. 1 (1918), pp. 115–43, first suggested an attack on the general problem of interpolation from the viewpoint of the theory of linear operations; certain of the results of his paper *Über Folgen linearer Operationen* (loc. cit.) can be applied to the problem of interpolation to continuous functions.

be a triangular sequence of real numbers such that $a \leqq t_{1n} < t_{2n} < \cdots < t_{nn} \leqq b$ $(n = 1, 2, \cdots)$, and let $\{h_{in}(t)\}$ $(i = 1, 2, \cdots, n; n = 1, 2, \cdots)$ be a set of continuous functions defined on (a, b), having the property $h_{in}(t_{kn}) = \delta_{ik}$, where δ is the Kronecker delta. The expression

$$H_n(t) = \sum_{i=1}^{n} x(t_{in}) h_{in}(t) \qquad (n = 1, 2, \cdots) \tag{7}$$

can then be considered a formula of interpolation to the function $x(t)$ in the interval (a, b). Let us suppose further that the fundamental functions $\{h_{in}(t)\}$ are positive[13] throughout (a, b). Theorem 2 enables us to conclude at once that the convergence of the sequence of positive linear functionals $U_n(x) \equiv H_n(\tau)$ $(n = 1, 2, \cdots)$ to the value $x(\tau)$ over P or T will imply the convergence of the given sequence over G_τ or $G_{\tau p}$. Hence the formula (7) converges, in this case, to the value of the function $x(t)$ at every point of continuity. Once more we have an illustration of the importance of the very general theorems of this section: the convergence properties of the formula (7) follow directly from Theorem 2 without any further considerations.

3. **Absolutely continuous limit-functionals.** We now pass to considerations which have important applications in the theory of mechanical quadratures. In this section attention will be confined to the set R of bounded functions integrable in the sense of Riemann. We introduce first some notations which materially simplify the discussion. Let J denote a finite set of disjoint intervals contained in (a, b), and J' its complement. With each x of R associate an element x_J of R as follows: $x_J(t) = x(t)$ or 0 according as t belongs to J or J'. With each operation $U(x)$ defined over R associate an operation $U^J(x)$ as follows: $U^J(x) = U(x_J)$. Clearly, if $U(x)$ is a linear functional over R, $U^J(x)$ is also; if $U(x)$ is positive over R, $U^J(x)$ has the same property. Evidently $U(x) = U^J(x) + U^{J'}(x)$ for every x of R.

The following is a generalization of theorems due to L. Fejér[14] and G. Pólya.[15]

THEOREM 5. *Let $\alpha(t)$ be absolutely continuous on (a, b). If a sequence $\{U_n(x)\}$ $(n = 1, 2, \cdots)$ of positive linear functionals defined over R converges over P to the functional $U(x) = \int_a^b x(t)\, d\alpha(t)$, then*[16] $\lim_{n\to\infty} U_n(x) = U(x)$ *for every x of R.*

[13] Interpolation formulas of this type have been studied by L. Fejér, *Über Interpolation*, Göttinger Nachrichten, vol. (1916), pp. 66–91, and D. Jackson, *The Theory of Approximation*, Colloquium Publications of American Mathematical Society, vol. 11 (1930), pp. 142–148. See also in this connection J. Shohat, *On interpolation*, Annals of Mathematics, vol. 34 (1933), pp. 130–146, and G. Szegö, *Interpolationspolynome*, Mathematische Zeitschrift, vol. 35 (1932), pp. 579–602.

[14] L. Fejér, *Mechanische Quadraturen mit positiven Côtesschen Zahlen*, Mathematische Zeitschrift, vol. 37 (1933), pp. 287–309.

[15] G. Pólya, *Über die Konvergenz von Quadraturverfahren*, Mathematische Zeitschrift, vol. 37 (1933), pp. 264–286.

[16] The integral is to be taken in the Riemann-Stieltjes sense. It is defined over R, since $\alpha(t)$ is absolutely continuous.

Under our hypothesis, and by Theorem 1, $\lim_{n\to\infty} U_n(x) = U(x)$ for every continuous function $x(t)$. Hence $U(x)$ is necessarily positive over C, being the limit, when $x \geqq \Theta$, of a sequence of non-negative numbers. We readily conclude that $\alpha(t)$ is monotone non-decreasing,[17] and from this it follows that $U(x)$ is positive over R.

Suppose now J is any finite set of disjoint closed intervals. We will show that $\lim_{n\to\infty} U_n^J(I) = U^J(I)$. Evidently, it will be sufficient to consider the case where J consists of a single closed interval (λ, μ), where $a < \lambda < \mu < b$. The proof can then be extended to the most general case by the use of the homogeneous-additive property of the linear functional. Let $x(t) = 1$ or 0 according as t belongs to J or J'. Then $U_n(x) = U_n^J(I)$ $(n = 1, 2, \cdots)$. We will show that $\lim_{n\to\infty} U_n(x)$ exists and has the value $U(x) = U^J(I)$. Let $\delta > 0$ be taken so small that $a \leqq \lambda - \delta$, $\lambda + \delta < \mu - \delta$ and $\mu + \delta \leqq b$. Define auxiliary functions $y(t)$ and $z(t)$, belonging to C, as follows:

$$y(t) = \begin{cases} 0 & a \leqq t \leqq \lambda - \delta \\ 1 & \lambda \leqq t \leqq \mu \\ 0 & \mu + \delta \leqq t \leqq b \\ \text{linear elsewhere,} \end{cases} \qquad z(t) = \begin{cases} 0 & a \leqq t \leqq \lambda \\ 1 & \lambda + \delta \leqq t \leqq \mu - \delta \\ 0 & \mu \leqq t \leqq b \\ \text{linear elsewhere.} \end{cases}$$

Now, $y \geqq x \geqq z$. Therefore, since the operations $\{U_n(x)\}$ and the operation $U(x)$ are positive over R,

$$U_n(y) \geqq U_n(x) \geqq U_n(z) \qquad (n = 1, 2, \cdots), \tag{8}$$

$$U(y) \geqq U(x) \geqq U(z). \tag{9}$$

From (8), $U(y) \geqq \overline{\lim_{n\to\infty}}\, U_n(x) \geqq U(z)$, for y and z are elements of C, whence $\lim_{n\to\infty} U_n(y) = U(y)$ and $\lim_{n\to\infty} U_n(z) = U(z)$. Finally, since

$$| U(y) - U(z) | = \left| \int_a^b [y(t) - z(t)]\, d\alpha(t) \right| \leqq 2V_\delta,$$

where V_δ represents the total variation of $\alpha(t)$ over the intervals $(\lambda - \delta, \lambda + \delta)$ and $(\mu - \delta, \mu + \delta)$, and V_δ is known to approach zero with δ, we conclude that $\lim_{n\to\infty} U_n(x)$ exists, and by (9) $\lim_{n\to\infty} U_n(x) = U(x)$. This proves our statement.

With this, we are in a position to establish the main theorem. Let x now be any element of R. Being given $\epsilon > 0$, by the criterion for Riemann integrability, it is possible to fit a net D on the interval (a, b) such that the sum of the lengths of those intervals of D is $< \epsilon$, in which the fluctuation of $x(t)$ is $\geqq \epsilon$. Let $\xi(t)$ be that continuous function which coincides with $x(t)$ at the end points of the intervals of D and is linear in the interior of each of these intervals. If we

[17] If we assume the contrary, there is no difficulty in constructing a non-negative continuous function $x(t)$ such that $\int_a^b x(t)\, d\alpha(t) < 0$.

denote by J the interval-set, throughout which the fluctuation of $x(t)$ is $\geqq \epsilon$, evidently $| x(t) - \xi(t) | < \epsilon$ for all values of t in the intervals J' complementary to J. We may take J to be closed. By our construction $m(J) < \epsilon$. Now

$$| U(x) - U_n(x) | \equiv | U_n(\xi - x) + U(x - \xi) + U(\xi) - U_n(\xi) |$$
$$\leqq | U_n(\xi - x) | + | U(x - \xi) | + | U(\xi) - U_n(\xi) |.$$

Using the fact that $| U_n^J |_R = U_n^J(I)$ (by Lemma 1), and since $\| x - \xi \| \leqq 2 \| x \|$, we conclude in succession,

$$| U_n(\xi - x) | \leqq | U_n^J(\xi - x) | + | U_n[(\xi - x)_{J'}] |$$
$$\leqq 2 \| x \| \, | U_n^J |_R + \epsilon | U_n |_R = 2 \| x \| \, U_n^J(I) + \epsilon \, U_n(I),$$
$$| U(x - \xi) | \leqq | U^J(x - \xi) | + | U[(x - \xi)_{J'}] |$$
$$\leqq 2 \| x \| \, | U^J |_R + \epsilon | U |_R = 2 \| x \| \, U^J(I) + \epsilon U(I);$$

and since ξ belonging to C implies $\lim_{n \to \infty} | U(\xi) - U_n(\xi) | = 0$, evidently

$$\overline{\lim_{n \to \infty}} \, | U(x) - U_n(x) | \leqq 4 \| x \| \, U^J(I) + 2\epsilon U(I), \tag{10}$$

for we have shown that $\lim_{n \to \infty} U_n^J(I) = U^J(I)$. Since $U^J(I) = \int_J d\alpha(t)$ represents simply the variation of the absolutely continuous function $\alpha(t)$ over the set J, and $m(J) < \epsilon$, it is evident that the quantity on the right hand side of (10) can be made arbitrarily small by choosing ϵ sufficiently small, and Theorem 5 is established.

Examination of the proof of Theorem 5 will show that the only property of the limit-functional $U(x)$ really needed in the demonstration (in addition to linearity) is the absolute continuity property which it has in common with the ordinary Riemann integral. This property may be characterized as follows: a linear functional $U(x)$, defined over R, will be said to be *absolutely continuous*, provided that for every x of R, $\epsilon > 0$ arbitrarily given implies the existence of a positive number $\delta = \delta(\epsilon, x)$ such that if J is any finite set of disjoint intervals contained in (a, b), and $m(J) < \delta$, then[18] $| U^J(x) | < \epsilon$. The question arises whether we cannot obtain greater generality by replacing $\int_a^b x(t) \, d\alpha(t)$ with a more general absolutely continuous limit-functional. The fact is, *every linear functional $U(x)$ defined over R which possesses the absolute continuity property above described is necessarily of the form $\int_a^b x(t) \, d\alpha(t)$, where $\alpha(t)$ is an absolutely continuous function*. This statement may be proved as follows. With the methods which Banach (loc. cit., p. 59) has used to establish the well-known theorem of

[18] In effect, we have called $U(x)$ absolutely continuous if, for fixed x, $U^J(x)$ is an absolutely continuous interval-set function $\Phi_x(J)$ in the sense of de La Vallée-Poussin, *Intégrales de Lebesgue*, Borel Collection, 1916, p. 57.

F. Riesz[19] on the general form of linear functionals defined over C-space, there is no difficulty in showing that every absolutely continuous linear functional $U(x)$ defined over R is of the form $\int_a^b x(t)d\alpha(t)$ ($\alpha(t)$ absolutely continuous) over C-space. We have to show that the given expression is the only extension of this functional to R-space which has the required property. For this we need

LEMMA 3. *If a linear functional $U(x)$ has the absolute continuity property over R, $|U^J|_R \to 0$ uniformly with $m(J)$ (i.e., $U(x)$ has this property uniformly over every bounded subset of R).*

We wish to show that $\epsilon > 0$ being given arbitrarily there exists a $\delta = \delta(\epsilon)$, such that if J is any finite set of disjoint intervals, and $m(J) < \delta$, then $|U^J(x)| < \epsilon$ for every x of R with $\|x\| \leqq 1$. Assume the contrary. Then there exists a $\rho > 0$, such that no matter how small $\delta > 0$ be taken, there is an $x = x_\delta$ (of norm $\leqq 1$) and a set $J = J_\delta$, with $m(J) < \delta$, such that $U^J(x) \geqq \rho$. That is, there exists a sequence of elements $\{x_n\}$ $(n = 1, 2, \cdots)$ of R of norm $\leqq 1$, and a sequence of sets J_n $(n = 1, 2, \cdots)$ with the property $\lim_{n\to\infty} m(J_n) = 0$, such that

$$U^{J_n}(x_n) \geqq \rho \qquad (n = 1, 2, \cdots). \tag{11}$$

Let N be a given positive integer, and let $\delta_k > 0$ $(k = 1, 2, \cdots)$ be so chosen that $|U^J(x_k)| < 1/2^k$ for every J such that $m(J) < \delta_k$. This is possible on account of the absolute continuity property of $U(x)$. Since $m(J_n)$ approaches zero with $1/n$, it is possible to select a set of indices $n = n_1, n_2, \cdots, n_N$ such that the part which J_{n_k} $(k = 1, 2, \cdots, N)$ has in common with the sets which follow (namely, $J_{n_{k+1}}, J_{n_{k+2}}, \cdots, J_{n_N}$) is of measure $< \delta_{n_k}$; let us call the sequence of sets obtained by removing this part $L_1, L_2, \cdots, L_N$. It follows that no two of the sets $\{L_k\}$ can have a point in common; each L_k is a subset of J_{n_k} and differs in measure from that of J_{n_k} by less than δ_{n_k}. Suppose now η is that element of R which is zero throughout the complement of $(L_1 + L_2 + \cdots + L_N)$, and in the set L_k $(k = 1, 2, \cdots, N)$ coincides with x_{n_k}. We have:

$$U(\eta) = \sum_{k=1}^N U^{L_k}(x_{n_k}) = \sum_{k=1}^N U^{J_{n_k}}(x_{n_k}) - \sum_{k=1}^N U^{J_{n_k}-L_k}(x_{n_k}),$$
$$\geqq N\rho - \sum_{k=1}^N \frac{1}{2^{n_k}} \geqq N\rho - 1,$$

in virtue of (11) and the fact that $|U^J(x_{n_k})| < 1/2^{n_k}$ for every J such that $m(J) < \delta_{n_k}$(in particular, $J = J_{n_k} - L_k$); since $\|\eta\| \leqq 1$, and N may be chosen arbitrarily large, this contradicts the boundedness of the linear functional $U(x)$, and the lemma is proved.

[19] F. Riesz, Annales de l'École Normale Supérieure, (3), vol. 31 (1914), pp. 9–14. This theorem has been extended to the class of bounded functions with at most discontinuities of the first kind by H. S. Kaltenborn, Bulletin of the American Mathematical Society, vol. 40(1934), pp. 702–708, and by T. H. Hildebrandt, Transactions of the American Mathematical Society, vol. 36(1934), pp. 868–875, to the class of all bounded measurable functions.

With this lemma, our original statement can easily be proved. Let x be any element of R. As we have seen in the proof of Theorem 5, there exists an element ξ of C, and a finite set J of disjoint intervals with $m(J) < \epsilon$, such that $| x(t) - \xi(t) | < \epsilon$ (t in $c(J) = J'$), $\| x - \xi \| \leqq 2 \| x \|$. Then, since $U(\xi) = \int_a^b \xi(t) d\alpha(t)$,

$$\left| U(x) - \int_a^b x(t)\, d\alpha(t) \right| = \left| U(x - \xi) - \int_a^b [x(t) - \xi(t)]\, d\alpha(t) \right|$$

$$\leqq | U(x - \xi) | + \left| \int_a^b [x(t) - \xi(t)]\, d\alpha(t) \right|$$

$$\leqq 2 \| x \| \, | U^J |_R + \epsilon \, | U |_R + 2 \| x \| \int_J d\alpha(t) + \epsilon V_\alpha ,$$

where $U(x - \xi) = U^J(x - \xi) + U^{J'}(x - \xi)$, etc., and where V_α denotes the total variation of $\alpha(t)$ over (a, b). Since, by Lemma 3, $| U^J |_R \to 0$ uniformly with $m(J)$, as $\epsilon \to 0$, and since $\alpha(t)$ is absolutely continuous, the right member of this inequality can be made arbitrarily small by choosing ϵ sufficiently small. Hence the left member must be zero, and our statement is proved.

As an application of Theorem 5, consider the formulas of mechanical quadratures of Gauss' type

$$\int_a^b x(t) p(t)\, dt = \sum_{i=1}^{n} H_{in} x(t_{in}) + R_n(x) \qquad (n = 1, 2, \cdots),$$

where $p(t)$ is non-negative and summable in (a, b), determined by the conditions

$$\text{(12)} \qquad R_n(t^k) \equiv 0 \qquad (k = 0, 1, 2, \cdots, 2n - 1;\ n = 1, 2, \cdots).$$

It can be shown that

$$H_{in} = \int_a^b \left[\frac{\Phi_n(t)}{(t - t_{in}) \Phi_n'(t_{in})} \right]^2 p(t)\, dt \qquad (i = 1, 2, \cdots, n;\ n = 1, 2, \cdots)$$

$$\left[\Phi_n(t) \equiv (t - t_{1n})(t - t_{2n}) \cdots (t - t_{nn}),\ \int_a^b \Phi_n \Phi_m p(t)\, dt = 0 \ (m \neq n) \right],$$

whence[20] $H_{in} > 0$ for all i and n. Putting $U_n(x) \equiv \sum_{i=1}^{n} H_{in} x(t_{in})$, $U(x) \equiv \int_a^b x(t)\, d\alpha(t)$, where $\alpha(t) \equiv \int_a^t p(t)\, dt$, and applying Theorem 5, we conclude that the formulas of Gauss' type converge for all bounded R-integrable functions $\{x(t)\}$, since convergence over P is assured by (12). The operations $\{U_n(x)\}$ are positive, since $H_{in} > 0$ ($i = 1, 2, \cdots, n;\ n = 1, 2, \cdots$).

[20] J. Shohat, *Théorie générale des polynomes orthogonaux de Tchebichef*, Mémorial des Sciences Mathématiques, LXVI (1934), p. 15.

In the paper referred to at the beginning of this section,[14] Fejér showed that *any* mechanical quadratures formulas of the type

$$\int_{-1}^{+1} x(t)\,dt = \sum_{i=1}^{n} H_{in}x(t_{in}) + R_n(x) \qquad (n = 1, 2, \cdots) \tag{13}$$

based on the Lagrange interpolation formula (i.e., $R_n(t^k) = 0$, $k = 0, 1, 2, \cdots, n - 1$) must converge for all R-integrable functions $x(t)$ if $H_{in} \geqq 0$. Pólya[15] has also given theorems from which the same conclusion can be derived. This result follows at once from Theorem 5.

4. **Non-functional positive operations. Uniform convergence.** We have shown, in the preceding pages, how the characterization of a certain type of linear functional as positive proves useful in discussing the convergence properties of sequences of such functionals; we shall now try to indicate how this characterization can be extended to include certain types of linear operations which are not necessarily functional. Consider, for example, an operation $U(x)$ defined over a vectorial subset E of G which associates with every x of E an element $y \equiv U(x)$ of some subset E_1 of G. E_1 is called the *range* or *contra-domain* of the operation $U(x)$. We make the following definition: *$U(x)$ is positive over E, provided $x \geqq \Theta$ implies $y \geqq \Theta$.* It is clear that we have here a direct extension of the definition of positiveness given previously for functionals. We have simply replaced the contra-domain r of real numbers by another normal vectorial space in which the symbol $\geqq$ is defined.

The statements made in §1 concerning the definition and properties of linear functionals defined over G apply without exception to linear operations with domain E and contra-domain E_1. In general, we use the same notations as in the previous case, though it should be noted that here the norm of $U(x)$ (as an element of G) must be written $\| U(x) \|$. An analogue to Lemma 1 follows at once.

Lemma 4. *If the linear operation $U(x)$, defined over a vectorial subset E of G which contains $x = I$, and with contra-domain lying in G, is positive over E, then $| U |_E = \| U(I) \|$.*

Exactly the same argument as that given in the proof of Lemma 1 will suffice to establish Lemma 4, if we note that in G, as well as with real numbers, $-x \leqq y \leqq x$ implies $\| y \| \leqq \| x \|$.

Similarly, Theorem 1 has its analogue, which we can state as follows:

Theorem 6. *Let E be a vectorial subset of G containing an element of positive lower bound in (a, b). If a sequence $\{U_n(x)\}$ $(n = 1, 2, \cdots)$ of positive linear operations defined over E, and with contra-domain lying in G, converges over a subset H dense in E to a linear operation $U(x)$ defined over E, with contra-domain in G, then*[21] $\lim_{n\to\infty} U_n(x) = U(x)$ *for every x of E.*

[21] Where previously we dealt with sequences of numbers, here it is a question of the convergence of a sequence of *functions* $\{U_n(x)\}$. This convergence is uniform, in view of the metric of G.

Again, the proof follows exactly the same lines as that of Theorem 1. We see without difficulty that the sequence converges for an element $\eta(t)$ of E such that $-\eta \leqq x \leqq \eta$ for all x of E with $\| x \| \leqq 1$. It follows that $| U_n |_E \leqq \| U_n(\eta) \|$ $(n = 1, 2, \cdots)$. This proves the theorem, for it shows that the set of norms $\{| U_n |_E\}$ is bounded. From this follows the convergence of the sequence to $U(x)$ at all points[22] of E.

As an application of Theorem 6, consider the integral of Fejér. Let

$$U_n(x) \equiv y_n(t) \equiv \frac{1}{2n\pi} \int_0^{2\pi} x(s) \left\{ \frac{\sin \frac{1}{2}n(s - t)}{\sin \frac{1}{2}(s - t)} \right\}^2 ds \qquad (n = 1, 2, \cdots). \tag{14}$$

The sequence (14) is a sequence of positive operations defined over C_p, $[(a, b) \equiv (0, 2\pi)]$, with contra-domain lying in C_p, which converges to the operation $U(x) = x(t)$ over T; by Theorem 6 the sequence of functions $\{y_n(t)\}$ converges uniformly over $(0, 2\pi)$ to the value $x(t)$ for every function $x(t)$ continuous in $(0, 2\pi)$ and of period 2π.

We can prove further that the convergence of the Fejér integral to the value $x(t)$ is uniform over any subinterval of $(0, 2\pi)$ in which $x(t)$ is continuous. Results of this kind can evidently be obtained for any singular integral or interpolation formula of positive character.

It is clear that the idea of positiveness in connection with linear operations can be extended to include operations defined over the most general abstract space of the normal vectorial type, with elements of any nature whatsoever, provided only some convenient meaning be assigned to the symbol $\geqq$ in domain and contra-domain. Whether such a classification of linear operations will be useful will depend to a great extent upon the type of space. The author hopes, at some future date, to extend the idea to the Hilbert spaces, with applications to the theory of integral equations in view.

5. **Coincidence of the formulas of mechanical quadratures of Gauss' type and of Tchebichef's type.** In §3 we called attention to the convergence properties of certain mechanical quadratures formulas of Gauss' type. In this section we consider a special problem connected with such formulas. Let $\psi(t)$ be a function bounded and non-decreasing, with infinitely many points of increase, over an interval (a, b) finite or infinite, and such that all moments $\alpha_n = \int_a^b t^n d\psi(t)$ $(n = 0, 1, 2, \cdots)$ exist, with $\alpha_0 > 0$. It is known that there exists an infinite sequence of orthogonal Tchebichef polynomials

$$\Phi_n(t) = \prod_{i=1}^{n} (t - t_{in}) \equiv t^n + p_{1n} t^{n-1} + \cdots + p_{nn},$$

[22] Banach, loc. cit., p. 79.

with all roots real, distinct, and lying in (a, b), determined by the conditions[23] $\int_a^b t^k\Phi_n(t)d\psi(t) = 0$ ($k = 0, 1, 2, \cdots, n-1; n = 1, 2, \cdots$). In order to avoid trivial equivalences, we assume $\psi(a) = 0$ and $\int_a^{a+h} d\psi(t) \int_{b-h}^b d\psi(t) > 0$ for $h > 0$. Under these conditions $\psi(t)$ will be called a characteristic function in (a, b). Two characteristic functions will be considered distinct only if they differ at a point of continuity.

Let $x(t)$ be such that $\int_a^b x(t)d\psi(t)$ exists in the Riemann-Stieltjes sense. Using the Lagrange interpolation polynomial which coincides with $x(t)$ at the points $\{t_{in}\}$ $(i = 1, 2, \cdots, n)$, we construct the so-called mechanical quadratures formula of Gauss' type

$$\int_a^b x(t)\,d\psi(t) = \sum_{i=1}^n H_{in}x(t_{in}) + R_n(x), \tag{15}$$

$$H_{in} = \int_a^b \frac{\Phi_n(t)\,d\psi(t)}{(t - t_{in})\,\Phi_n'(t_{in})}$$

having the property $R_n(t^k) = 0$ ($k = 0, 1, \cdots, 2n-1$). If (a, b) is finite, it is well known that for $\psi(t) \equiv \int_a^t \frac{dt}{\pi\sqrt{(t-a)(b-t)}}$, which gives rise to the so-called trigonometric polynomials $\cos n \arccos [(a + b - 2t)/(a - b)]$, the Côtes numbers H_{in} are equal for each n: $H_{in} = H_n$ ($i = 1, 2, \cdots, n; n = 1, 2, \cdots$), and the Gauss formula (15) is therefore, in this special case, at the same time of Tchebichef's type.[24] The question naturally arises whether the two types of formulas coincide in any other case. We propose to show that *the two formulas cannot coincide in any other case.*[25]

Our method is to use the hypothesis of coincidence to derive conditions on the moments $\{\alpha_k\}$ ($k = 0, 1, 2, \cdots$) of $\psi(t)$, and by this means to show that $\psi(t)$ is uniquely determined in the class of admissible characteristic functions. In

[23] For the basic theory used throughout this section, see J. Shohat, *Théorie générale des polynomes orthogonaux de Tchebichef*, loc. cit.

[24] Mechanical quadratures formulas characterized by the property of possessing equal Côtes numbers were first discussed by P. Tchebichef, Journal de Mathématiques, vol. 19 (1874), pp. 19–34.

[25] The above result was obtained by the author late in 1933 and presented to the American Mathematical Society in November, 1934. I learned recently from Professor Shohat that Professor M. Krawtchouk in June 1934 presented to the All Russian Mathematical Congress in Leningrad a similar theorem, but only for the special case $\psi(x) \equiv \int p(x)dx$. Cf. M. Krawtchouk, *Sur une question algébrique dans le problème des moments*, Journal de l'Institut Mathématique de l'Académie des Sciences de l'Ukraine, vol. 2 (1934), pp. 87–92; in Ukrainian.

fact, since $[\Phi_n(s) - \Phi_n(t)]/(s - t)$ is a polynomial of degree $n - 1$ in t, if the formula (15) is of the Gauss-Tchebichef type, we must have[26]

$$\begin{aligned}\int_a^b \frac{\Phi_n(s) - \Phi_n(t)}{s - t}\, d\psi(t) &= H_n \sum_{i=1}^{n} \frac{\Phi_n(s) - \Phi_n(t_{in})}{s - t_{in}} \\ &= H_n \Phi_n(s) \sum_{i=1}^{n} \frac{1}{s - t_{in}} = H_n \Phi_n'(s).\end{aligned} \tag{16}$$

Replacing $\Phi_n(s)$, $\Phi_n(t)$, $\Phi_n'(s)$ in (16) by their explicit expressions in terms of the coefficients $\{p_{in}\}$ of $\Phi_n(t)$ and performing the indicated operations, we conclude immediately that for $n = 1, 2, \cdots$,

$$\begin{cases} n\alpha_1 + \alpha_0 p_{1n} = 0, \\ n\alpha_2 + np_{1n}\alpha_1 + 2p_{2n}\alpha_0 = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ n\alpha_{n-1} + np_{1n}\alpha_{n-2} + \cdots + (n - 1)\, p_{n-1,n}\alpha_0 = 0. \end{cases} \tag{17}$$

To these relations we may add

$$\alpha_n + p_{1n}\alpha_{n-1} + \cdots + p_{nn}\alpha_0 = 0 \qquad (n \geqq 1), \tag{18}$$

$$\alpha_{n+1} + p_{1n}\alpha_n + \cdots + p_{nn}\alpha_1 = 0 \qquad (n \geqq 2), \tag{19}$$

in view of the orthogonality properties of $\Phi_n(t)$. The equations (17), together with (18), give the coefficients $\{p_{kn}\}$ $(k = 1, 2, \cdots, n)$ in terms of $\alpha_0, \alpha_1, \cdots, \alpha_n$, since the determinant of the system is clearly different from zero; therefore, in virtue of (19), α_{n+1} $(n \geqq 2)$ is determined by $\alpha_0, \alpha_1, \cdots, \alpha_n$. That is, the moments $\{\alpha_k\}$ $(k = 0, 1, 2, \cdots)$ are completely determined by α_0, α_1, α_2, through a recurrence formula which is independent of (a, b) and $\psi(t)$.

We have only one more step. There is no loss of generality in assuming $\alpha_0 = 1$; by the Schwarz inequality, it follows that

$$\alpha_1^2 \leqq \alpha_2. \tag{20}$$

Moreover, (20) is precisely the necessary and sufficient condition for the existence of real numbers c, d such that $\alpha_1 = \frac{1}{2}(c + d)$, $\alpha_2 = (c - d)^2/8 + (c + d)^2/4$. Inasmuch as all the moments $\{\alpha_k\}$ are determined by α_1 and α_2, it follows that the set $\{\alpha_k\}$ is identical with the set of moments associated with the characteristic function $\bar{\psi}(t) \equiv \int_c^t \frac{dt}{\pi\sqrt{(t - c)(d - t)}}$ in the finite interval (c, d), for the moments of this latter function satisfy the same relations as the set $\{\alpha_k\}$, and

[26] This interesting formula shows that, when coincidence takes place, the numerators $\{\Omega_n(t)\}$ of the successive convergents of the associated continued fraction (Shohat, loc. cit., p. 12) are simply the derivatives of the polynomials themselves. J. Sherman, *Stieltjes continued fractions*, Transactions of the American Mathematical Society, vol. 35 (1933), pp. 64–87, p. 81 has pointed out that this cannot occur with the classical polynomials of Jacobi, Laguerre and Hermite except in the trigonometric case.

$\int_c^d d\bar{\psi}(t) = 1$, $\int_c^d t d\bar{\psi}(t) = \frac{1}{2}(c + d)$, $\int_c^d t^2 d\bar{\psi}(t) = (c - d)^2/8 + (c + d)^2/4$, as can easily be verified. But *the set of moments* $\{\alpha_k\}$ $(k = 0, 1, 2, \cdots)$ *associated with a characteristic function* $\psi(t)$ *in a finite interval* (a, b) *cannot be generated by a different characteristic function in the same or any other interval, finite or infinite.*[27] It follows that coincidence cannot take place in the infinite interval, and that in the finite interval (a, b) the known solution $(c, d) \equiv (a, b)$,

$$\psi(t) \equiv \int_a^t \frac{dt}{\pi\sqrt{(t - a)(b - t)}}$$

is unique.

UNIVERSITY OF PENNSYLVANIA.

[27] As Prof. Shohat pointed out to the author, this fact is an immediate consequence of theorems due to Stieltjes and Carleman on the determinateness of the moment problem. Cf. J. Shohat, loc. cit., p. 7.

THE ZEROS OF JACOBI AND RELATED POLYNOMIALS

By C. Eugene Buell

Introduction

1. **Definitions.** The ultraspherical polynomials of degree n, $P_n^{(\lambda)}(\cos \vartheta)$, are defined as polynomials not vanishing identically for which the differential equation

$$y'' + \{(n + \lambda)^2 + \lambda(1 - \lambda) \sin^{-2}\vartheta\}y = 0 \tag{1}$$

has the solution $y = \sin^\lambda\vartheta \cdot P_n^{(\lambda)}(\cos \vartheta)$. It will also be convenient to consider the generating function of these polynomials normalized in a proper way, namely,

$$(1 - 2w \cos \vartheta + w^2)^{-\lambda} = \sum_{n=0}^{\infty} P_n^{(\lambda)}(\cos \vartheta) \cdot w^n. \tag{2}$$

The Jacobi polynomials of degree n, $P_n^{(\alpha, \beta)}(\cos \vartheta)$, are defined as polynomials not vanishing identically for which the differential equation

$$y'' + \left\{\left(n + \frac{\alpha + \beta + 1}{2}\right)^2 + \frac{\frac{1}{4} - \alpha^2}{4 \sin^2 \theta/2} + \frac{\frac{1}{4} - \beta^2}{4 \cos^2 \theta/2}\right\} y = 0 \tag{3}$$

has the solution $y = [\sin (\vartheta/2)]^{\alpha+\frac{1}{2}} [\cos (\vartheta/2)]^{\beta+\frac{1}{2}} \cdot P_n^{(\alpha, \beta)}(\cos \vartheta)$.

The Jacobi polynomials reduce to the ultraspherical polynomials if $\alpha = \beta = \lambda - \frac{1}{2}$. The ultraspherical polynomials reduce to the Legendre polynomials if $\lambda = \frac{1}{2}$. Concerning further properties of these polynomials, we refer[1] to **[5]** and **[8]**.

2. **Previous estimates.** For $\lambda > -\frac{1}{2}$ all of the zeros of the ultraspherical polynomials are real. Let ϑ_k denote the k-th zero in increasing order, $0 < \vartheta_k < \pi$. The following estimates for ϑ_k have been given:

(1) Bruns **[2]** for Legendre polynomials:

$$\text{(A)} \qquad \frac{k - \frac{1}{2}}{n + \frac{1}{2}}\pi < \vartheta_k < \frac{k}{n + \frac{1}{2}}\pi \qquad (k = 1, 2, \cdots, n).$$

(2) A. Markoff **[4]** and Stieltjes **[7]** for Legendre polynomials:

$$\text{(B)} \qquad \frac{k - \frac{1}{2}}{n}\pi < \vartheta_k < \frac{k}{n + 1}\pi \qquad \left(k = 1, 2, \cdots, \left[\frac{n}{2}\right]\right).$$

Received December 2, 1935.

[1] Numbers in bold face type refer to the bibliography at the end of this paper.

(3) Szegö [9] for $0 < \lambda < 1$:

$$(C_1)\qquad \frac{k+\lambda-1}{n+\lambda}\pi < \vartheta_k < \frac{k}{n+\lambda}\pi \qquad (k = 1, 2, \cdots, n),$$

$$(C_2)\qquad \frac{k+\dfrac{\lambda-1}{2}}{n+\lambda}\pi < \vartheta_k < \frac{k}{n+1}\pi \qquad \left(k = 1, 2, \cdots, \left[\frac{n}{2}\right]\right),$$

$$(C_3)\qquad \frac{j_k}{\sqrt{(n+\lambda)^2+\lambda(1-\lambda)c}} < \vartheta_k < \frac{j_k}{n+\lambda} \qquad \left(k = 1, 2, \cdots, \left[\frac{n+1}{2}\right]\right),$$

where j_k denotes the k-th positive zero of the Bessel function of order $\lambda - \frac{1}{2}$ and $c = 1 - (2/\pi)^2$.

The purpose of this paper is to obtain estimates for the zeros of ultraspherical polynomials for some values of $\lambda < 0$ and for $\lambda > 1$, and to obtain analogous estimates for the zeros of Jacobi polynomials.

3. **Preliminary theorems.** In the following, we shall consider the ultraspherical and Jacobi polynomials (apart from a factor) as solutions of an ordinary second order differential equation of the form $y'' + \varphi(x)y = 0$. This will be compared with another differential equation of the same form, the zeros of whose solution will be considered as known. The following theorems due essentially to Sturm will form the basis of this comparison.[2]

THEOREM 1. *Let $f(x)$ and $F(x)$ be two continuous functions on $a < x \leqq b$ such that $f(x) \leqq F(x)$ and $f(x) \not\equiv F(x)$. Consider the ordinary differential equations $y'' + f(x)y = 0$, $Y'' + F(x)Y = 0$, and let $y(x)$ be a solution of the first and $Y(x)$ be a solution of the second. Suppose that*

$$(1)\qquad y(x) > 0 \text{ on } a < x < b \text{ and } y(b) = 0,$$

$$(2)\qquad \lim_{x\to a+0}\{y'(x)Y(x) - y(x)Y'(x)\} \geqq 0.$$

Then either $Y(x) \equiv 0$ or there exists a point ξ, $a < \xi < b$, such that $Y(\xi) < 0$.

If (2) is replaced by

$$(2')\qquad \lim_{x\to a+0}\{y'(x)Y(x) - y(x)Y'(x)\} \leqq 0,$$

then either $Y(x) \equiv 0$ or there exists a point ξ, $a < \xi < b$, such that $Y(\xi) > 0$.

If we replace (2) by

$$(2'')\qquad \lim_{x\to a+0}\{y'(x)Y(x) - y(x)Y'(x)\} = 0,$$

then either $Y(x) \equiv 0$ or there exists a point ξ, $a < \xi < b$, such that $Y(x)$ has a variation of sign in ξ.

[2] Cf. Szegö, [9].

If we omit (2) and replace (1) by

(1′) $\quad y(x) > 0 \quad$ on $\quad a < x < b \quad$ and $\quad y(a) = y(b) = 0,$

then either $Y(x) \equiv 0$ or there exists a point ξ, $a < \xi < b$, such that $Y(x)$ has again a variation of sign in ξ. (The ordinary Sturm theorem.)

THEOREM 2. *Consider $y'' + \varphi(x)y = 0$, where $\varphi(x)$ is continuous and monotonically decreasing (increasing) in the strict sense on $a < x < b$. Suppose that a solution $y(x)$ has the zeros $\alpha, \beta, \gamma, \cdots$ (at least three and arranged in increasing order) on (a, b). Then $\beta - \alpha < \gamma - \beta < \cdots$ ($\beta - \alpha > \gamma - \beta > \cdots$).*

Remark. For the inequality $\beta - \alpha < \gamma - \beta$ ($\beta - \alpha > \gamma - \beta$) we need only the fact that $\varphi(x') > \varphi(x'')$ ($\varphi(x') < \varphi(x'')$) for $\alpha < x' < \beta$ and $\beta < x'' < \gamma$.

I. Ultraspherical polynomials for $\lambda > 1$

1. **Trigonometric comparison.** The ultraspherical polynomials can be characterized by the differential equation

(1) $$y'' + \{(n + \lambda)^2 + \lambda(1 - \lambda) \sin^{-2}\vartheta\}y = 0$$

with the solution $y = \sin^\lambda \vartheta \cdot P_n^{(\lambda)}(\cos \vartheta) \equiv \vartheta^\lambda \cdot f(\vartheta)$, where $f(\vartheta) = a_0 + a_1\vartheta + \cdots$, $a_0 \neq 0$. We compare this with the differential equation

(2) $$u'' + (n + \lambda)^2 u = 0$$

with the solution $u = \sin(n + \lambda)\vartheta = \vartheta g(\vartheta) = \vartheta(b_0 + b_1\vartheta + \cdots)$, where $b_0 \neq 0$. Since for $\lambda > 1$, $\lambda(1 - \lambda) < 0$, because of Sturm's oscillation theorem between two consecutive zeros of $y(\vartheta)$ there lies at least one zero of $u(\vartheta)$. Hence

(3) $$\vartheta_k - \vartheta_{k-1} > \frac{\pi}{n + \lambda} \qquad (k = 1, 2, \cdots, n).$$

This inequality holds for $k = 1$, $\vartheta_0 = 0$, since

$$\lim_{\vartheta \to +0} \{y'(\vartheta)u(\vartheta) - y(\vartheta)u'(\vartheta)\} = 0.$$

Adding (3) for successive values of k, we obtain

(4) $$\vartheta_k > \frac{k}{n + \lambda}\pi.$$

Now $\vartheta_k + \vartheta_{n+1-k} = \pi$, whence

(5) $$\vartheta_k = \pi - \vartheta_{n+1-k} < \frac{k + \lambda - 1}{n + \lambda}\pi.$$

The combination of (4) and (5) gives the estimate for the k-th zero of $P_n^{(\lambda)}(\cos \vartheta)$:

(6) $$\frac{k}{n + \lambda}\pi < \vartheta_k < \frac{k + \lambda - 1}{n + \lambda}\pi \qquad (k = 1, 2, \cdots, n).$$

This corresponds to the estimate (C_1) given by Szegö for $0 < \lambda < 1$ (Introduction 2); it is particularly like that of Bruns for the Legendre polynomials. It is evident that if $\lambda = 1$, $\vartheta_k = k\pi/(n + 1)$.

2. **The analogue of Szegö's estimate** (C_2). From Theorem 2 of the Introduction we have $\vartheta_1 - 0 > \vartheta_2 - \vartheta_1 > \cdots > \vartheta_{[n/2]+1} - \vartheta_{[n/2]}$, since the coefficient of y in (1) is monotonically increasing in $(0, \pi/2)$; furthermore, for n even, it fulfills the conditions given in the remark to Theorem 2 with $\alpha = \vartheta_{[n/2]-1}$, $\beta = \vartheta_{[n/2]}$, $\gamma = \vartheta_{[n/2]+1}$. Thus the polygonal line joining the points (k, ϑ_k), $0 \leqq k \leqq [n/2] + 1$, is concave upward and hence takes on its minima at its end points. Consider now $\vartheta_k' = \vartheta_k - kc_1 - c_2$. The line joining the points (k, ϑ_k') is also concave upward. We want to determine c_1 and c_2 so that $\vartheta_k' \geqq 0$ at each end point and so that the inequality $\vartheta_k \geqq kc_1 + c_2$ will be "better" than (4).

For $k = 0$, $\vartheta_0' = \vartheta_0 - c_2 = -c_2$; we put $c_2 = 0$. Moreover for n odd,

$$\vartheta_{(n+1)/2}' = \vartheta_{(n+1)/2} - \frac{n+1}{2} c_1 = \frac{\pi}{2} - \frac{n+1}{2} c_1.$$

If we put this equal to zero, $c_1 = \pi/(n+1)$. For n even we obtain

$$\vartheta_{n/2}' = \vartheta_{n/2} - \frac{\pi}{n+1} \cdot \frac{n}{2}, \qquad \vartheta_{1+n/2}' = \vartheta_{1+n/2} - \frac{\pi}{n+1}\left(\frac{n}{2} + 1\right).$$

Adding, we obtain

$$\vartheta_{n/2}' + \vartheta_{1+n/2}' = \vartheta_{n/2} + \vartheta_{1+n/2} - \pi = 0.$$

Now since the polygonal line joining the points (k, ϑ_k') is concave, we cannot have $\vartheta_{n/2}' < 0, \vartheta_{1+n/2}' > 0$ or $\vartheta_{n/2}' = \vartheta_{n/2+1}' = 0$. Hence $\vartheta_{n/2}' > 0, \vartheta_{1+n/2}' < 0$. Thus we have for both n odd and n even

$$\vartheta_k' = \vartheta_k - k\pi/(n+1) \geqq 0 \qquad (k = 1, 2, \cdots, [(n+1)/2]),$$

the equality holding for n odd and $k = (n+1)/2$. The lower estimate

$$\vartheta_k > k\pi/(n+1) \tag{7}$$

is better than (4).

Consider now

$$\vartheta_k'' = \vartheta_k - \frac{k+c}{n+\lambda}\pi.$$

Then $\vartheta_k'' - \vartheta_{k-1}'' = \vartheta_k - \vartheta_{k-1} - \pi/(n+\lambda)$. From (3), we have $\vartheta_k'' - \vartheta_{k-1}'' > 0$. Thus the sequence $\{\vartheta_k''\}$ is monotonically increasing and takes on its maximum at its right end point. We want to determine c so that this maximum will be less than or equal to zero.

For n odd, $\vartheta_{(n+1)/2} = \pi/2$ so that

$$\vartheta_{(n+1)/2}'' = \frac{\pi}{2} - \frac{\frac{n+1}{2} + c}{n+\lambda}\pi.$$

If we set this equal to zero, then $2c = \lambda - 1$. For n even, it must be verified that

$$\vartheta''_{n/2} = \vartheta_{n/2} - \frac{\frac{n}{2} + \frac{\lambda - 1}{2}}{n + \lambda}\,\pi < 0.$$

Now

$$\vartheta_{1+n/2} - \vartheta_{n/2} = \pi - 2\vartheta_{n/2} > \pi/(n + \lambda),$$

whence the statement follows. The upper estimate

$$\vartheta_k < \frac{k + \frac{\lambda - 1}{2}}{n + \lambda}\,\pi \qquad \left(k = 1, 2, \cdots, \left[\frac{n}{2}\right]\right) \tag{8}$$

thus obtained is better than (5). The combination of (7) and (8) gives the estimate

$$\frac{k}{n + 1}\,\pi < \vartheta_k < \frac{k + \frac{\lambda - 1}{2}}{n + \lambda}\,\pi \quad \left(k = 1, 2, \cdots, \left[\frac{n}{2}\right]\right), \tag{9}$$

in which, compared with (6), both the upper and lower estimates are improved.

3. **Bessel comparison.** The inequalities

$$\vartheta^{-2} < \sin^{-2}\vartheta \leqq \vartheta^{-2} + c, \qquad c = 1 - (2/\pi)^2, \qquad 0 < \vartheta \leqq \pi/2 \tag{10}$$

lead from (1) to the comparison equations

$$u'' + \{(n + \lambda)^2 + \lambda(1 - \lambda)\vartheta^{-2}\}u = 0, \tag{11}$$

$$v'' + \{(n + \lambda)^2 + \lambda(1 - \lambda)c + \lambda(1 - \lambda)\vartheta^{-2}\}v = 0, \tag{12}$$

which have the solutions

$$u = \sqrt{\vartheta}\cdot J_\gamma\{(n + \lambda)\vartheta\} = \vartheta^\lambda(a_0 + a_1\vartheta + \cdots), \qquad \gamma = \lambda - \tfrac{1}{2},\ a_0 \neq 0, \tag{11'}$$

and

$$v = \sqrt{\vartheta}\cdot J_\gamma\{[(n + \lambda)^2 + \lambda(1 - \lambda)c]^{1/2}\vartheta\} = \vartheta^\lambda(b_0 + b_1\vartheta + \cdots), \qquad \gamma = \lambda - \tfrac{1}{2},\ b_0 \neq 0, \tag{12'}$$

respectively. Between the differential equations (1), (11) and (12) we have relations similar to those in Theorem 1 for $0 < \vartheta < \pi/2$. Further, it is easily seen that

$$\lim_{\vartheta \to +0} \{u'(\vartheta)y(\vartheta) - u(\vartheta)y'(\vartheta)\} = 0$$

and

$$\lim_{\vartheta \to +0} \{y'(\vartheta)v(\vartheta) - y(\vartheta)v'(\vartheta)\} = 0.$$

Thus between consecutive zeros of $y(\vartheta)$ there is at least one zero of $u(\vartheta)$ and between consecutive zeros of $v(\vartheta)$ there is at least one zero of $y(\vartheta)$. Hence, denoting the zeros of $J_\gamma(\vartheta)$ by $j_1, j_2, j_3, \cdots$ in increasing order, we have the estimate

$$\frac{j_k}{n+\lambda} < \vartheta_k < \frac{j_k}{[(n+\lambda)^2 + \lambda(1-\lambda)c]^{\frac{1}{2}}} \cdot \quad \left(k = 1, 2, \cdots, \left[\frac{n+1}{2}\right]\right). \tag{13}$$

These inequalities correspond to the estimate (C_3) of Szegö (Introduction 2).

4. **An estimate for j_k.** Now consider k fixed and $n = 2k - 1, k = (n+1)/2$ Then from (13)

$$\frac{j_k}{2k-1+\lambda} < \frac{\pi}{2} < \frac{j_k}{\sqrt{(2k-1+\lambda)^2 + \lambda(1-\lambda)c}},$$

from which

$$\left(k - \frac{1-\lambda}{2}\right)\pi\sqrt{1 + \frac{\lambda(1-\lambda)c}{(2k-1+\lambda)^2}} < j_k < \left(k - \frac{1-\lambda}{2}\right)\pi,$$

whence, writing $\gamma = \lambda - \frac{1}{2}$, we have

$$\left(k - \frac{1}{4} + \frac{\gamma}{2}\right)\pi\sqrt{1 + \frac{(\frac{1}{4} - \gamma^2)c}{(2k - \frac{1}{2} + \gamma)^2}} < j_k < \left(k - \frac{1}{4} + \frac{\gamma}{2}\right)\pi.$$

We thus have

$$j_k = \left(k - \frac{1}{4} + \frac{\gamma}{2}\right)\pi\cdot\left\{1 + O\left(\frac{1}{k^2}\right)\right\}. \tag{14}$$

II. Estimates derived from Stieltjes' integral representation[3]

1. **Stieltjes' integral representation.** Using the generating function of the ultraspherical polynomials (see eq. 2, Introduction), Stieltjes derived the following important representation:[4]

$$\begin{aligned} P_n^{(\lambda)}(\cos\vartheta) &= 2\pi^{-1}\sin\pi\lambda(2\sin\vartheta)^{-\lambda}\,\Re e^{i[(n+\lambda)\vartheta - \pi\lambda/2]} \\ &\qquad \int_0^1 r^{-\lambda}(1-r)^{n+2\lambda-1}(1-kr)^{-\lambda}\,dr, \end{aligned} \tag{1}$$

where

$$k = \tfrac{1}{2}(1 - i\cot\vartheta). \tag{2}$$

[3] In a letter to Prof. Szegö (January 7, 1935) Prof. Fejér gave a proof of the estimate (C_2) of Szegö in the special case of Legendre polynomials ($\lambda = \frac{1}{2}$) based on Stieltjes integral representation. The results of this part were derived after I had seen this letter through the kindness of Prof. Szegö. In a letter (February 10, 1935) Fejér derives (C_2) generally from (1) for $0 < \lambda < 1$. Finally in a paper which should appear in the *Monatshefte für Mathematik und Physik* he obtains an upper estimate for ϑ_k which is for $0 < \lambda < \frac{1}{2}$ better than that of (C_2) as well as better than the upper bound in (5).

[4] [1], II, 122 (in a letter dated December 19, 1890). Cf. also Szegö [8], p. 57.

Here the first two factors of the integrand are real and positive; the third factor is equal to unity for $r = 0$. This complex factor may be written for $0 < \vartheta < \pi/2$ in the form $(1 - kr)^{-\lambda} = \rho(r)^{-\lambda}e^{-i\lambda\psi(r)}$, where $\rho(r) > 0$, $\psi(r)$ real and $\rho(0) = 0$, $\psi(0) = 0$. Now

$$\psi(r) = \tan^{-1}\left(\cot\vartheta\cdot\frac{r}{2-r}\right),$$

so if

$$\frac{A_n(\vartheta)}{B_n(\vartheta)} = \int_0^1 r^{-\lambda}(1-r)^{n+2\lambda-1}\rho(r)^{-\lambda}\frac{\cos}{\sin}\left\{\lambda\tan^{-1}\left(\cot\vartheta\cdot\frac{r}{2-r}\right)\right\}dr, \tag{3}$$

then

$$\begin{aligned} P_n^{(\lambda)}(\cos\vartheta) &= 2\pi^{-1}\sin\pi\lambda(2\sin\vartheta)^{-\lambda} \\ &\quad \{A_n(\vartheta)\cos[(n+\lambda)\vartheta - \pi\lambda/2] + B_n(\vartheta)\sin[(n+\lambda)\vartheta - \pi\lambda/2]\}. \end{aligned} \tag{4}$$

2. **An estimate for** $0 < \lambda < \frac{1}{2}$. Consider now the following values for $(n+\lambda)\vartheta - \pi\lambda/2$: $k\pi$, $(k\pi - \pi/4)$, $(k\pi - \pi/2)$, $(k\pi - 3\pi/4)$. The corresponding values of ϑ are respectively

$$a_k = \frac{k + \frac{\lambda}{2}}{n+\lambda}\pi, \quad b_k = \frac{k - \frac{1-2\lambda}{4}}{n+\lambda}\pi, \quad c_k = \frac{k - \frac{1-\lambda}{2}}{n+\lambda}\pi, \quad d_k = \frac{k - \frac{3-2\lambda}{4}}{n+\lambda}\pi,$$

while the corresponding values for $\frac{1}{2}\pi(2\sin\vartheta)^{\lambda}(\sin\pi\lambda)^{-1}P_n^{(\lambda)}(\cos\vartheta)$ are respectively

$$(-1)^k A_n(a_k), \qquad (-1)^k 2^{-\frac{1}{2}}\{A_n(b_k) - B_n(b_k)\}, \qquad (-1)^{k+1}B_n(c_k),$$
$$(-1)^{k+1}2^{-\frac{1}{2}}\{A_n(d_k) + B_n(d_k)\}.$$

Now since $0 < \lambda < \frac{1}{2}$ and $0 < \vartheta < \pi/2$, $0 < \lambda\tan^{-1}[\cot\vartheta\cdot r/(2-r)] < \pi/4$, so that $A_n(\vartheta)$, $B_n(\vartheta)$, and $A_n(\vartheta) - B_n(\vartheta)$ are all positive. Therefore $P_n^{(\lambda)}(\cos\vartheta)$ changes sign between b_k and c_k and hence must have a zero there. This gives the following estimate:

$$\frac{k - \frac{1-\lambda}{2}}{n+\lambda}\pi < \vartheta_k < \frac{k - \frac{1-2\lambda}{4}}{n+\lambda}\pi \qquad \left(k = 1, 2, \cdots, \left[\frac{n}{2}\right]\right). \tag{5}$$

The upper estimate given by Szegö [see Introduction, (C_2)] is $k\pi/(n+1)$. For values of k less than $\frac{\frac{1}{2}-\lambda}{1-\lambda}\cdot\frac{n+1}{2}$, the estimate above is better.

3. **An estimate for** $-\frac{1}{2} < \lambda < 0$. For $-\frac{1}{2} < \lambda < 0$ and $0 < \vartheta < \pi/2$

$$-\frac{\pi}{4} < \lambda\tan^{-1}\left(\cot\vartheta\cdot\frac{r}{2-r}\right) < 0,$$

so that $A_n(\vartheta)$ and $A_n(\vartheta) + B_n(\vartheta)$ are positive, while $B_n(\vartheta)$ is negative. Thus $P_n^{(\lambda)}(\cos\vartheta)$ must have a zero between c_k and d_k. This gives the following estimate:

$$(6) \qquad \frac{k - \dfrac{3-2\lambda}{4}}{n+\lambda}\pi < \vartheta_k < \frac{k - \dfrac{1-\lambda}{2}}{n+\lambda}\pi \qquad \left(k = 1, 2, \cdots, \left[\frac{n}{2}\right]\right).$$

Remark. (5) and (6) also hold for n odd and $k = (n+1)/2$, when $=$ is put instead of $<$ in the lower estimate in (5) and similarly in the upper estimate in (6).

III. Jacobi polynomials

1. Trigonometric comparison. For the Jacobi polynomials we have the differential equation

$$y'' + \left\{\left(n + \frac{\alpha+\beta+1}{2}\right)^2 + \frac{\frac{1}{4}-\alpha^2}{4\sin^2\vartheta/2} + \frac{\frac{1}{4}-\beta^2}{4\cos^2\vartheta/2}\right\} y = 0$$

with the solution $y = (\sin\vartheta/2)^{\alpha+\frac{1}{2}}(\cos\vartheta/2)^{\beta+\frac{1}{2}}\cdot P_n^{(\alpha,\beta)}(\cos\vartheta) = \vartheta^{\alpha+\frac{1}{2}}\cdot f(\vartheta)$, where $f(\vartheta) = a_0 + a_1\vartheta + \cdots$, $a_0 \neq 0$. We compare this with the differential equation

$$u'' + \left(n + \frac{\alpha+\beta+1}{2}\right)^2 u = 0$$

with the solution $u = \sin\left(n + \dfrac{\alpha+\beta+1}{2}\right)\vartheta = \vartheta\cdot g(\vartheta)$, where $g(\vartheta) = b_0 + b_1\vartheta + \cdots$, $b_0 \neq 0$. For $\alpha, \beta > -1$ all of the zeros of $P_n^{(\alpha,\beta)}(\cos\vartheta)$ are real. Let them be denoted in increasing order by $\vartheta_k^{(\alpha,\beta)}$, $0 < \vartheta_k^{(\alpha,\beta)} < \pi$; $k = 1, 2, \cdots, n$. We discuss the position of $\vartheta_k^{(\alpha,\beta)}$ for $\alpha, \beta > -\frac{1}{2}$.

Case A: $\alpha^2 < \frac{1}{4}$, $\beta^2 < \frac{1}{4}$. In this case

$$(1) \quad \left(n + \frac{\alpha+\beta+1}{2}\right)^2 + \frac{\frac{1}{4}-\alpha^2}{4\sin^2\vartheta/2} + \frac{\frac{1}{4}-\beta^2}{4\cos^2\vartheta/2} > \left(n + \frac{\alpha+\beta+1}{2}\right)^2,$$

so that between consecutive positive zeros of $u(\vartheta)$ there lies at least one zero of $y(\vartheta)$. Further,

$$\lim_{\vartheta\to+0}\{u'y - y'u\} = 0,$$

since $\alpha + \frac{1}{2} > 0$. We see, therefore, that at least one zero of $P_n^{(\alpha,\beta)}(\cos\vartheta)$ lies in each interval

$$\left(\frac{k-1}{n + \dfrac{\alpha+\beta+1}{2}}\pi, \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\pi\right) \quad (k = 1, 2, \cdots, n).$$

Further, since the length of each of these intervals is less than π/n, we must have exactly one zero in each interval. From this the following estimates result:

$$(2) \qquad \frac{k-1}{n + \dfrac{\alpha+\beta+1}{2}}\pi < \vartheta_k^{(\alpha,\beta)} < \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\pi \qquad (k = 1, 2, \cdots, n).$$

The above argument may be carried through for $P_n^{(\beta,\alpha)}(\cos\vartheta)$, since $\beta + \frac{1}{2} > 0$. Thus $\vartheta_k^{(\beta,\alpha)}$ also satisfies (2). Now[5]

$$\vartheta_k^{(\alpha,\beta)} + \vartheta_{n+1-k}^{(\beta,\alpha)} = \pi, \tag{3}$$

so that

$$\vartheta_k^{(\alpha,\beta)} = \pi - \vartheta_{n+1-k}^{(\beta,\alpha)} > \frac{k + \dfrac{\alpha+\beta-1}{2}}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi.$$

Since in this case $-1 < \dfrac{\alpha+\beta-1}{2} < 0$, the lower estimate of (2) has been improved. The combination of this lower estimate and the upper estimate from (2) gives the following:

$$\frac{k + \dfrac{\alpha+\beta-1}{2}}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi < \vartheta_k^{(\alpha,\beta)} < \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi \qquad (k = 1, 2, \cdots, n). \tag{4}$$

For $\alpha = \beta = \lambda - \frac{1}{2}$ we obtain Szegö's estimates (C_2) (Introduction 2).

Remark. If the comparison equation

$$u'' + \left\{\left(n + \frac{\alpha+\beta+1}{2}\right)^2 + \frac{\frac{1}{2} - \alpha^2 - \beta^2}{4}\right\} u = 0$$

is used, the upper estimate is improved slightly. Thus

$$\frac{k + \dfrac{\alpha+\beta-1}{2}}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi < \vartheta_k^{(\alpha,\beta)} < \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi \times \left[1 + \frac{\frac{1}{2} - \alpha^2 - \beta^2}{4\left(n + \dfrac{\alpha+\beta+1}{2}\right)^2}\right]^{-\frac{1}{2}} \qquad (k = 1, 2, \cdots, n). \tag{4'}$$

Case B: $\alpha > \frac{1}{2}$, $\beta > \frac{1}{2}$. In this case, the inequality (1) must be reversed. Again

$$\lim_{\vartheta \to +0} (y'u - u'y) = 0,$$

since $\alpha > \frac{1}{2}$, so that

$$\vartheta_1^{(\alpha,\beta)} > \frac{\pi}{n + \dfrac{\alpha+\beta+1}{2}}.$$

[5] Cf. for instance [8], p. 4, (3).

Applying Sturm's oscillation theorem and the reversed inequality (1), we have

$$\vartheta_k^{(\alpha,\beta)} - \vartheta_{k-1}^{(\alpha,\beta)} > \frac{\pi}{n + \dfrac{\alpha+\beta+1}{2}}.$$

Adding these inequalities for successive values of k, we obtain

$$\vartheta_k^{(\alpha,\beta)} > \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi.$$

The relation $\vartheta_k^{(\alpha,\beta)} + \vartheta_{n+1-k}^{(\beta,\alpha)} = \pi$ gives

$$\vartheta_k^{(\alpha,\beta)} = \pi - \vartheta_{n+1-k}^{(\beta,\alpha)} < \pi - \frac{n+1-k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi = \frac{k + \dfrac{\alpha+\beta-1}{2}}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi.$$

The last two inequalities give an estimate that is analogous to (4), namely:

$$(5) \qquad \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi < \vartheta_k^{(\alpha,\beta)} < \frac{k + \dfrac{\alpha+\beta-1}{2}}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi \qquad (k = 1, 2, \cdots, n).$$

For $\alpha = \beta = \lambda - \frac{1}{2}$ we obtain I, (6).

Case C: $\alpha^2 < \frac{1}{4}$, $\beta > \frac{1}{2}$ *and Case* D: $\alpha > \frac{1}{2}$, $\beta^2 < \frac{1}{4}$. For these cases[6] there is a value of ϑ, say ϑ', $0 < \vartheta' < \pi$, for which

$$\frac{\frac{1}{4} - \alpha^2}{\sin^2 \vartheta/2} + \frac{\frac{1}{4} - \beta^2}{\cos^2 \vartheta/2} = 0.$$

Then in the interval $(0, \vartheta')$ we have

$$\vartheta_k^{(\alpha,\beta)} - \vartheta_{k-1}^{(\alpha,\beta)} \lessgtr \frac{\pi}{n + \dfrac{\alpha+\beta+1}{2}}$$

in cases C and D respectively from the same considerations as above. Adding for successive values of k,

$$\begin{matrix}(6)\\(7)\end{matrix} \qquad \vartheta_k^{(\alpha,\beta)} \lessgtr \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi$$

in cases C and D, respectively, provided that $0 < \vartheta_k^{(\alpha,\beta)} < \vartheta'(\alpha, \beta)$.

In order to obtain estimates for the interval (ϑ', π) we use (3). If we interchange α and β, cases C and D are interchanged and ϑ' becomes $\pi - \vartheta'$; i.e.,

[6] We always assume tacitly $\alpha, \beta > -1$.

$\vartheta'(\beta, \alpha) = \pi - \vartheta'(\alpha, \beta)$. Consequently

$$\vartheta_k^{(\beta,\alpha)} \gtrless \frac{k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi$$

in cases C and D, respectively, provided that $0 < \vartheta_k^{(\beta,\alpha)} < \vartheta'(\beta, \alpha)$. If now $\vartheta'(\alpha, \beta) \leqq \vartheta_k^{(\alpha,\beta)} < \pi$, we have $0 < \vartheta_{n+1-k}^{(\beta,\alpha)} \leqq \vartheta'(\beta, \alpha)$. Therefore

$$\vartheta_k^{(\alpha,\beta)} = \pi - \vartheta_{n+1-k}^{(\beta,\alpha)} \lessgtr \pi - \frac{n+1-k}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi$$

or

$$\begin{matrix}(8)\\(9)\end{matrix} \qquad \vartheta_k^{(\alpha,\beta)} \lessgtr \frac{k + \dfrac{\alpha+\beta-1}{2}}{n + \dfrac{\alpha+\beta+1}{2}}\,\pi \qquad [\vartheta_k^{(\alpha,\beta)} \text{ in } (\vartheta', \pi)]$$

in cases C and D, respectively. Thus the process that gave both upper and lower estimates in cases A and B on the interval $(0, \pi)$ only yields an upper estimate in case C and a lower estimate in case D on the intervals $(0, \vartheta')$ and (ϑ', π).

2. **Bessel comparison.** The following representation of the coefficient in the equation of the Jacobi polynomials is convenient for obtaining estimates of their zeros by means of those of the Bessel functions:

$$\frac{\frac{1}{4} - \alpha^2}{4\sin^2 \vartheta/2} + \frac{\frac{1}{4} - \beta^2}{4\cos^2 \vartheta/2} = \frac{\frac{1}{2} - \alpha^2 - \beta^2}{2\sin^2 \vartheta} + \frac{\beta^2 - \alpha^2}{2} \cdot \frac{\cos \vartheta}{\sin^2 \vartheta}.$$

We have

$$\vartheta^{-2} < \sin^{-2} \vartheta < \vartheta^{-2} + c, \qquad 0 < \vartheta \leq \varphi < \pi,$$

where $c = \sin^{-2} \varphi - \varphi^{-2}$ and $1 - \vartheta^2/2 < \cos \vartheta < 1$, whence for $\varphi = \pi/2$

$$(10) \qquad \vartheta^{-2} - \tfrac{1}{2} < \frac{\cos \vartheta}{\sin^2 \vartheta} < \vartheta^{-2} + c.$$

For brevity, we put

$$(11) \quad \Theta(\alpha, \beta; \vartheta) = \left(n + \frac{\alpha+\beta+1}{2}\right)^2 + \frac{\frac{1}{2} - \alpha^2 - \beta^2}{2\sin^2 \vartheta} + \frac{\beta^2 - \alpha^2}{2} \cdot \frac{\cos \vartheta}{\sin^2 \vartheta}.$$

There are four cases depending on the signs of the expressions $\frac{1}{2} - \alpha^2 - \beta^2$, $\beta^2 - \alpha^2$.

Case I: $\frac{1}{2} - \alpha^2 - \beta^2 > 0, \beta^2 - \alpha^2 > 0$. Using (10) we have[7]

[7] We again assume throughout $\alpha, \beta > -1$.

$$\left(n + \frac{\alpha + \beta + 1}{2}\right)^2 - \frac{\beta^2 - \alpha^2}{4} + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2} < \Theta(\alpha, \beta; \vartheta)$$
$$< \left(n + \frac{\alpha + \beta + 1}{2}\right)^2 + (\tfrac{1}{4} - \alpha^2)c + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2}.$$

Denoting by A^2 and B^2 the terms independent of ϑ, $A > 0$, $B > 0$, we obtain[8]

$$A^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2} < \Theta(\alpha, \beta; \vartheta) < B^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2}.$$

Case II: $\frac{1}{2} - \alpha^2 - \beta^2 > 0, \beta^2 - \alpha^2 < 0, \alpha > -\frac{1}{2}$. As in case I we have

$$\left(n + \frac{\alpha + \beta + 1}{2}\right)^2 + \frac{\beta^2 - \alpha^2}{2}c + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2} < \Theta(\alpha, \beta; \vartheta)$$
$$< \left(n + \frac{\alpha + \beta + 1}{2}\right)^2 + \frac{(\frac{1}{2} - \alpha^2 - \beta^2)c}{2} - \frac{\beta^2 - \alpha^2}{4} + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2},$$

which may be written in the form

$$C^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2} < \Theta(\alpha, \beta; \vartheta) < D^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2}.$$

Case III: $\frac{1}{2} - \alpha^2 - \beta^2 < 0, \beta^2 - \alpha^2 < 0$. In this case, the constant terms are the same as in case I and appear interchanged. Thus

$$B^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2} < \Theta(\alpha, \beta; \vartheta) < A^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2}.$$

Case IV: $\frac{1}{2} - \alpha^2 - \beta^2 < 0, \beta^2 - \alpha^2 > 0, \alpha > -\frac{1}{2}$. In this case the constant terms are the same as in Case II and appear interchanged. Thus

$$D^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2} < \Theta(\alpha, \beta; \vartheta) < C^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2}.$$

The comparison equation in all four cases is of the form

$$u'' + \left\{A^2 + \frac{\frac{1}{4} - \alpha^2}{\vartheta^2}\right\}u = 0, \qquad A > 0,$$

with the solution $u = \vartheta^{\frac{1}{2}} \cdot J_\alpha(A\vartheta) = \vartheta^{\alpha + \frac{1}{2}} \cdot g(\vartheta)$, where $g(\vartheta) = a_0 + a_1\vartheta + \cdots$, $a_0 \neq 0$. Further, for all cases,

$$\lim_{\vartheta \to +0} (y'u - u'y) = -\lim_{\vartheta \to +0} (u'y - y'u) = 0,$$

since $\alpha + \frac{1}{2} > 0$. The resulting estimates are then

Case I: $j_k/B < \vartheta_k^{(\alpha,\beta)} < j_k/A$;
Case II: $j_k/D < \vartheta_k^{(\alpha,\beta)} < j_k/C$;
Case III: $j_k/A < \vartheta_k^{(\alpha,\beta)} < j_k/B$;
Case IV: $j_k/C < \vartheta_k^{(\alpha,\beta)} < j_k/D$,

[8] We take n so large that $A^2 > 0$, $B^2 > 0$.

where j_k denotes the k-th zero of the Bessel function $J_\alpha(\vartheta)$ of order α and $k = 1, 2, \cdots, k'$, where k' is such that

$$\vartheta_{k'}^{(\alpha,\beta)} < \varphi < \vartheta_{k'+1}^{(\alpha,\beta)},$$

and n is large enough so that all the expressions A, B, C, and D are real.

There is no difficulty in the discussion of the cases in which either of the two expressions $\frac{1}{4} - \alpha^2 - \beta^2, \beta^2 - \alpha^2$, is zero.

Bibliography

[1] B. Baillaud and H. Bourget, *Correspondance d'Hermite et de Stieltjes*, vols. I, II, Paris, 1905.

[2] H. Bruns, *Zur Theorie der Kugelfunktionen*, Journal für die reine und angewandte Mathematik, vol. 90 (1881), pp. 322–328.

[3] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*, vol. 1, Berlin, 1924.

[4] A. Markoff, *Sur les racines de certaines équations*, Mathematische Annalen, vol. 27 (1886), pp. 177–182.

[5] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. II, Berlin, 1925.

[6] J. Shohat, *Théorie générale des polynomes orthogonaux de Tschebichef*, Mémorial des Sciences Mathématiques, vol. 66, Paris, 1934.

[7] Th. Stieltjes, *Sur les racines de l'équation* $X_n = 0$, Acta Mathematica, vol. 9 (1887), pp. 385–400.

[8] G. Szegö, *Asymptotische Entwicklungen der Jacobischen Polynome*, Königsberger Gelehrte Gesellschaft, 1933.

[9] G. Szegö, *Inequalities for the zeros of Legendre polynomials and related functions*, Trans. Amer. Math. Soc., vol. 39 (1936), pp. 1–17.

[10] E. Hille, *Über die Nullstellen der Hermiteschen Polynome*, Jahresbericht Deutsch. Math. Ver., vol. 44 (1934), pp. 162–165.

[11] J. Shohat, *On a certain formula of mechanical quadratures with non-equidistant ordinates*, Trans. Amer. Math. Soc., vol. 31 (1929), pp. 448–463.

Washington University.

CLASSES OF MAXIMUM NUMBERS ASSOCIATED WITH CERTAIN SYMMETRIC EQUATIONS IN *n* RECIPROCALS. III

By H. A. Simmons and W. E. Block

1. **Introduction.** By extending considerably the methods used by Simmons in the first[1] paper I, and by Stelford and Simmons in the second[2] paper II, we shall obtain results that include as special cases all theorems of I, II (cf. the definition of *remarkable properties* in this section and[3] Theorems 5, 8, 9 and 12). We shall explain in more detail what we do in this paper after we recall from I a few definitions that we use here.

If a solution $x \equiv (x_1, \cdots, x_n)$ of any given equation with which we deal is obtained by Kellogg's process[4] of minimizing the variables $x_1, \cdots, x_{n-1}$ in this order, one at a time, we shall denote it by w and call it the *Kellogg solution* of the given equation. For the equations that we consider the Kellogg solution is (except in §14) one in positive integers. It always belongs to the general class of solutions that we admit, namely, that in which $x_1, \cdots, x_{n-1}$ are positive integers and $x_1 \leqq x_2 \leqq \cdots \leqq x_n$. These solutions include all positive integral solutions and are called *E-solutions* (for extended solutions, beyond those in positive integers). Thus, for a very simple example, the Kellogg solution of $x_1^{-1} + x_2^{-1} = 2/7$ is $x = w = (4, 28)$ and its E-solutions are (4, 28), (5, 35/3), (6, 42/5), and (7, 7). From Theorem 2, p. 887, of I, we know that 28 $(=w_2)$ is the largest number that exists in any E-solution of the given equation and that 28 appears in no E-solution of this equation except w. Furthermore, if $P(x_1, x_2) \equiv P(x)$ is any symmetric polynomial in x_1, x_2 with no negative coefficient, and if $P(x)$ is not a mere constant, Theorem 3 of I contains the following statement as a very special fact: if $x = X$ is any E-solution of the equation $x_1^{-1} + x_2^{-1} = 2/7$ other than its Kellogg solution w, then $P(X) < P(w)$.

Where nothing is said to the contrary, we adopt generally the definitions and notation of I, II. Thus $P(x)$ stands for a polynomial of the type defined above except that $P(x)$ contains n variables instead of 2; and with $i \geqq 0$ and j equal to

Received August 12, 1935; presented to the American Mathematical Society, April 19, 1935. The results of Parts 1, 2, 3 of this paper are due to Simmons; those of Part 4, chiefly to Block, a student at Northwestern University.

[1] Cf. Trans. Amer. Math. Soc., vol. 34 (1932), pp. 876–907.

[2] Cf. Bulletin Amer. Math. Soc., vol. 40 (1934), pp. 884–894.

[3] A theorem has the same number as the section which contains it.

[4] Concerning Kellogg's diophantine problem and extensions of it, cf. O. D. Kellogg, American Mathematical Monthly, vol. 28 (1921), p. 300; D. R. Curtiss, ibid., vol. 29 (1922), pp. 381–387; and Tanzô Takenouchi, Proceedings of the Physico-Mathematical Society of Japan, (3), vol. 3, pp. 78–92.

integers, we let $\Sigma_{i,j}(x)$ stand for the j-th elementary symmetric function of the i variables $x_1, \cdots, x_i$ with the customary agreement that

$$\Sigma_{i,j}(x) \begin{cases} \equiv 0 \text{ when } i < j \text{ and also when } j < 0 , \\ \equiv 1 \text{ when } j = 0 . \end{cases}$$

For brevity, when we have that the Kellogg solution w of a given equation (e) contains (in the sense described above) the individual maximum number that exists in any E-solution of (e) and maximizes $P(X)$, where X is any E-solution of (e), so that $P(X) < P(w)$ if $X \neq w$, we shall say that *w has the remarkable properties relative to* (e). The former (latter) of these properties will be referred to as the *first* (*second*) remarkable property.

In I and II we established the remarkable properties for the Kellogg solution w, of each of the following equations in the x_i:

$$\Sigma_{n,r}(1/x) + \lambda\Sigma_{n,r+1}(1/x) = b/a , \qquad a \equiv [(c+1)\,b - 1] , \tag{1}$$

$$\Sigma_{n,r}(1/x) + \Sigma_{n,r+1}(1/x) + \cdots + \Sigma_{n,s}(1/x) = b/a , \tag{1.1}$$

where r, s, n are integers such that $1 \leqq r < s \leqq n$, λ is any integer $\geqq 0$, and b, c are any positive integers. In II we generalized the results in question for certain cases in which a is not of the form $(c + 1)\,b - 1$. The extensions which we make here can be understood from inspection of the equations that we treat here as we treated (1) and (1.1) in I, II, namely,[5]

$$\Sigma_{n,r}(1/x) + \Sigma_{n,n}(1/x) = b/a \qquad (n \geqq r+2) , \tag{1.2}$$

$$\begin{aligned} &\Sigma_{n,r}(1/x) + \lambda_{r+1}\Sigma_{n,r+1}(1/x) + \lambda_{r+2}\Sigma_{n,r+2}(1/x) + \cdots + \lambda_s\Sigma_{n,s}(1/x) = b/a \\ &\qquad\qquad\qquad\qquad\qquad\qquad (n \geqq s > r+1) , \end{aligned} \tag{1.3}$$

where the λ_t ($t = r + 1, \cdots, s$) are integers $\geqq 0$ such that

$$\begin{aligned} &\lambda_{i+1}\lambda_{j+2} - \lambda_{i+2}\lambda_{j+1} \geqq 0 \qquad (i = r-1, \cdots, s-1; j = r-2, \cdots, i-1) \\ &\lambda_{r-1} = \lambda_{s+1} = 0 , \quad \lambda_r = 1 ; \end{aligned} \tag{1.4}$$

$$\begin{aligned} &\Sigma_{n,r}(1/x) + \lambda_{r+1}\Sigma_{n,r+1}(1/x) + \lambda_{r+2}\Sigma_{n,r+2}(1/x) + \cdots + \lambda_s\Sigma_{n,s}(1/x) \\ &\qquad + \Sigma_{n,n}(1/x) = b/a \qquad (n \geqq s + 2 > n + 2) , \end{aligned} \tag{1.5}$$

where[6] the λ_t are as in (1.4).

[5] Were we to use $\Sigma_{n,r}(1/x) + \lambda\Sigma_{n,n}(1/x)$ in the place of the left member of (1.2), we would need to prove in §4 the following inequality instead of (4.14):

$$(1^s)^2 \geqq (1^{s+1})(1^{s-1}) + \lambda(1^{s-1}).$$

We have not been able to establish this inequality when λ is a positive integer restricted only by the new equation (1.2) in question. It is for a similar reason that we do not use a positive integral parameter λ as a multiplier of $\Sigma_{n,n}(1/x)$ in (1.5).

[6] To allow the case $n = (s + 1)$ here would duplicate a case of (1.3).

In (1.2), (1.3) and (1.5) a can be as it is defined for (1) for every set of values of n, r, s and the λ's that we admit, and for some choices of these numbers extra content is given to our theory by allowing a the generality that it has in II (cf. Theorem 8). One should note here that there are infinitely many sets of positive λ's for which (1.4) holds.[7] For example, with $s = (r + 2)$, (1.4) is satisfied if $\lambda_{r+1}^2 \geqq \lambda_{r+2} > 0$; with $s = (r + 3)$, (1.4) is true if $\lambda_{r+1}^2 \geqq \lambda_{r+2}$, $\lambda_{r+1}\lambda_{r+2} \geqq \lambda_{r+3} > 0$, and $\lambda_{r+2}^2 \geqq \lambda_{r+1}\lambda_{r+3}$. It is also to be observed that the second subscripts of the Σ's in (1.3) are consecutive and that the only cases where we do not require these subscripts to be such are exhibited in (1.2) and (1.5). We have tried, without success, to find for our theory some modification that would enable us to establish the remarkable properties for the Kellogg solution of the equation

$$\Sigma_{n,1}(1/x) + \Sigma_{n,3}(1/x) = b/a\,,$$

where a is as it is in (1) and n is any integer > 3. The difficulty is that we are unable to establish the analog of (2.3) below [or of (4.3) and (4.4) in the cases of (1.2) for which $r > 1$; or of (6.3) and (6.4) in our treatment of (1.3)]. For this reason, we have not been able to apply our method with success to any general equation of the form[8]

$$\Sigma_{n,r}(1/x) + \Sigma_{n,s}(1/x) = b/a \qquad (n > s > r + 1)\,.$$

New features of our procedure. The nature of our present modification of the procedure of I and II is described fairly well in the following four statements.

(1) In the case $r = 1$ of (1.2), the individual maximum number and the class of maximum numbers are identified in *two different processes*; the class of maximum numbers for this case is identified by transforming the elements of an arbitrary E-solution $X \neq w$ into their Kellogg correspondents in the *reverse* order, from X_n to X_1 (cf. §§2, 3).

(2) We use inequalities here that require extra detail (beyond that of I, II) relative to the sizes of the X_i in an E-solution [cf. (4.12), for example].

(3) The inequalities between the partition symbols in §§4, 11 involve extra terms [cf. the last term, (1^{s-1}), in (4.14) and the right member of (11.16)] and call for a variation of corresponding procedure[9] in II.

(4) The observations that (7) and (11) are equivalent to (7.1) and (11.1), respectively, are new and they are important because (7.1) and (11.1) are analogs[10] here of the key-inequality of I, namely, (46) of I.

The discussion from §2 to the end of this paper is divided into four parts as follows: Part 1 deals with (1.2), the case of (1.5) in which $\lambda_t = 0$ $(t = r + 1, \cdots, s)$, §§2 to 4 (inclusive); Part 2 with (1.3), §§5 to 8; Part 3 with the case of (1.5) in

[7] Inequality (1.4) implies that if $\lambda_r, \ldots, \lambda_k$ $(r \leqq k < s - 1)$ are positive and if $\lambda_{k+1} = 0$, then $\lambda_t = 0$ $(t = k + 1, \cdots, s)$, as can be readily proved.

[8] We deal with a special equation of this form in §13.

[9] Cf. p. 888 of II.

[10] The case $r = 1$ is admitted in (7.1) and (11.1), whereas it was excluded in (46) of I.

which $\lambda_t > 0$ (cf. footnote 8), §§9 to 12; Part 4 with problems that we have considered in trying to extend our theory and with an application of this theory to the convergence of certain types of series, §§13 to 16. The order in which we present our results (of Parts 1, 2, 3) is that in which we have obtained them[11] and is one of increasing difficulty. Part 3 depends upon (and naturally follows) Part 2 in the same way that Part 1 here depends upon I [cf. our proof of (4.3) below].

Part 1. The remarkable properties of the Kellogg solution of (1.2)

2. **The maximum number in any E-solution of a special equation.** We consider now the case $r = 1$ of (1.2),

$$\Sigma_{n,1}(1/x) + \Sigma_{n,n}(1/x) = b/a\,, \qquad a \equiv [(c+1)b - 1]\,. \tag{2}$$

The Kellogg solution of (2) is [cf. I, p. 886, equations (23)] $x = w$, where

$$\begin{aligned} &w_1 = c + 1\,, \qquad w_{p+1} = aw_1 \cdots w_p + 1 \qquad (p = 1, \cdots, n-2)\,,\\ &\qquad\qquad w_n = a(w_1 \cdots w_{n-1} + 1)\,. \end{aligned} \tag{2.1}$$

We wish to prove now that the w of (2.1) has the first remarkable property relative to equation (2).

Proof. From I we know that if $X_{1\cdots(n-1)}$ is any set of $n-1$ positive integers such that $\Sigma_{n-1,1}(1/X) < b/a$, then $X_{1\cdots(n-1)}$ is a set σ [cf. 14 of I] and

$$\Sigma_{p,1}(1/X) \leqq \Sigma_{p,1}(1/w) \qquad (p = 1, \cdots, n-1)\,. \tag{2.2}$$

Suppose now that X is an E-solution of (2), so that (2.2) holds. Further, let X be different from w and suppose that $X_n \geqq w_n$. We shall reach a contradiction. Under our hypothesis X contains at least two elements that differ from their corresponding elements of w, one element of class A and one of class B (cf. p. 891 of I) and, according to equation (31) of I, the inequality sign in (2.2) holds for $p = (n-1)$. Further, since $X_n \geqq w_n$,

$$\Sigma_{n,1}(1/X) < \Sigma_{n,1}(1/w)\,. \tag{2.3}$$

From (2.3) and the fact that X is a solution of (2), it follows that

$$X_1 \cdots X_n < w_1 \cdots w_n\,. \tag{2.4}$$

However, by solving (2) for x_n, we find that

$$X_n = a(X_1 \cdots X_{n-1} + 1)[bX_1 \cdots X_{n-1} - a\Sigma_{n-1,n-1}(X)]^{-1}\,,$$

so that

$$X_n \leqq a(X_1 \cdots X_{n-1} + 1)\,. \tag{2.5}$$

Since $w_n = a(w_1 \cdots w_{n-1} + 1)$, (2.4) and (2.5) are inconsistent with our hypothesis that $X_n \geqq w_n$. Therefore $X_n < w_n$.

[11] In the beginning of this work our main object was to overcome the difficulty that is explained in §22 of I.

3. **The class of maximum numbers relative to** (2). We wish to prove next that the w of (2.1) has the second remarkable property relative to (2).

Proof. We consider the elements of any E-solution $X \neq w$ in the order X_n, $X_{n-1}, \cdots, X_1$. Since $X \neq w$, X contains at least one element of each of the classes A and B, and we let X_{θ_1}, $X_{{}_1\theta}$ be the first elements in $X_{n \cdots 1}$ of class A, B, respectively. Our transformation from X to X' is an analog of (33) of I (cf. the transformation defined below by setting $f = X$ in (3)); then if $X' \neq w$, we repeat the transformation to pass from X' to X''; etc. Using the notation of p. 898 of I, we define our transformation from $f_{n \cdots 1}$ to $f'_{n \cdots 1}$ to be t_1 or t_2:

$$
(3) \quad
\begin{aligned}
&(t_1)\, f'_p = f_p \,(p \neq \theta_1, {}_1\theta)\,, \qquad f'_{\theta_1} = w_{\theta_1}\,, \\
&\qquad\qquad \Sigma_{n,1}\left(\frac{1}{f'}\right) + \Sigma_{n,n}\left(\frac{1}{f'}\right) = \Sigma_{n,1}\left(\frac{1}{f}\right) + \Sigma_{n,n}\left(\frac{1}{f}\right), \\
&(t_2)\, f'_p = f_p \,(p \neq \theta_1, {}_1\theta)\,, \qquad f'_{{}_1\theta} = w_{{}_1\theta}\,, \\
&\qquad\qquad \Sigma_{n,1}\left(\frac{1}{f'}\right) + \Sigma_{n,n}\left(\frac{1}{f'}\right) = \Sigma_{n,1}\left(\frac{1}{f}\right) + \Sigma_{n,n}\left(\frac{1}{f}\right),
\end{aligned}
$$

according as t_1 defines $f'_{{}_1\theta}$ to be not greater than $w_{{}_1\theta}$ or greater than $w_{{}_1\theta}$, respectively. We shall show that every transformation which we employ in passing from X to w is such that[12]

$$
(3.1) \qquad f_{\theta_1} f_{{}_1\theta} < f'_{\theta_1} f'_{{}_1\theta}\,, \qquad (f_{\theta_1})^t + (f_{{}_1\theta})^t < (f'_{\theta_1})^t + (f'_{{}_1\theta})^t,
$$

where t is any positive integer; the second remarkable property will then follow from Lemma 3 of I. That (3.1) holds for the transformation from X to X' follows from Lemma 1a of I and the fact that X is an E-solution of (2) different from w so that $X_n = X_{{}_1\theta} < w_n$ (cf. §2). We shall be assured of the validity of (3.1) generally if we can show that in any intermediate set $f_{n \cdots 1} = X^{(i)}$ the first element which differs from its correspondent in w is $f_{{}_1\theta}$, of class $B^{(i)}$. Suppose this were not the case. Then we should have

$$
(3.2) \quad f_p = X_p \;(p = 1, \cdots, \theta_1 - 1), \quad f_{\theta_1} > w_{\theta_1}, \quad f_p = w_p \;(p = \theta_1 + 1, \cdots, n).
$$

Now (3.2) implies that $f_{1 \cdots (\theta_1 - 1)} \neq w_{1 \cdots (\theta_1 - 1)}$. Hence from (3.2) and (31) of I, it follows that

$$
(3.3) \qquad \Sigma_{\theta_1 - 1, 1}(1/f) < \Sigma_{\theta_1 - 1, 1}(1/w).
$$

Then (3.2) and (3.3) imply the inequalities

$$
(3.4) \qquad \Sigma_{\theta_1, 1}(1/f) < \Sigma_{\theta_1, 1}(1/w),
$$

$$
(3.5) \qquad \Sigma_{n, 1}(1/f) < \Sigma_{n, 1}(1/w).
$$

[12] That $f_{\theta_1} f_{{}_1\theta} \neq f'_{\theta_1} f_{{}_1\theta}$ follows from the second footnote on p. 892 of I.

Hence we could replace f_{θ_1} by w_{θ_1} in (3.4) and (3.5) and obtain true inequalities. Indeed we could make the substitution $f'_{\theta_1} = w_{\theta_1}$, with $f'_p = f_p$ ($p \neq \theta_1, {}_1\theta$, where ${}_1\theta = n$), and select $f'_n > w_n$ in such a way that the new set f' would satisfy (2). This set f' would be an E-solution of (2), and the inequality $f'_n > w_n$ would contradict the theorem proved in §2, namely, that w_n has the first remarkable property relative to (1.2).

4. **The cases** $r > 1$ **of** (1.2). We wish to establish the remarkable properties for the Kellogg solution w of (1.2) in the cases $r = 2, \cdots, n-2$; the case $r = n - 1$ of (1.2) was treated in I. The Kellogg solution $x = w$ of (1.2) is defined by the following equations

$$
\begin{gathered}
w_p = 1 \quad (p = 1, \cdots, r-1), \quad w_r = c + 1, \\
w_{p+1} = a\Sigma_{p,p-r+1}(w) + 1 \quad (p = r, \cdots, n-2), \quad w_n = a[\Sigma_{n-1,n-r}(w) + 1].
\end{gathered} \tag{4}
$$

We define the left member of (1.2) to be $\varphi_n(1/x)$, and write generally

$$
\varphi_p(1/x) \equiv \Sigma_{p,r}(1/x) + \Sigma_{p,n}(1/x) \qquad (r \leqq p \leqq n), \tag{4.1}
$$

so that $\varphi_p(1/x) = \Sigma_{p,r}(1/x)$ for $r \leqq p < n$.

From (4.1) it is clear that $\varphi_n(1/x)$ may be written in the form

$$
\begin{gathered}
\varphi_n(1/x) = \Sigma_{n-1,r}(1/x) + x_n^{-1}\psi(1/x), \\
\psi(1/x) \equiv \Sigma_{n-1,r-1}(1/x) + \Sigma_{n-1,n-1}(1/x).
\end{gathered} \tag{4.2}
$$

From (4.2) one observes that in order to establish the first remarkable property for w, it suffices to show that if X is any E-solution of (1.2) except its Kellogg solution w, then the following relations are true [cf. (30) and (32) of I]:

$$
\Sigma_{n-1,r}(1/X) \leqq \Sigma_{n-1,r}(1/w), \tag{4.3}
$$

$$
\psi(1/X) < \psi(1/w). \tag{4.4}
$$

The class of maximum numbers in which we are interested (cf. §1) will also be identified if we can show that every transformation that we make in passing from X to w (by one or more transformations) accords with (3.1) (cf. Lemma 3 of I). We shall prove this to be the case after we make two more definitions.

Definition of transformation. We now consider the elements of our solution $X \neq w$, and more generally of f [cf. (3)] in the order from the first to the last: $f = f_{1 \cdots n}$. If $f \equiv X^{(\alpha)}$ contains at least one element of each of the classes $A^{(\alpha)}, B^{(\alpha)}$, the transformation from f to f' is defined by t_3 or t_4,

$$
\begin{aligned}
&(t_3)\ f'_p = f_p \quad (p \neq \theta_1, {}_1\theta), \qquad f'_{\theta_1} = w_{\theta_1}, \qquad \varphi_n(1/f') = \varphi_n(1/f), \\
&(t_4)\ f'_p = f_p \quad (p \neq \theta_1, {}_1\theta), \qquad f'_{{}_1\theta} = w_{{}_1\theta}, \qquad \varphi_n(1/f') = \varphi_n(1/f),
\end{aligned} \tag{4.5}
$$

according as t_3 defines $f'_{{}_1\theta}$ to be not greater than $w_{{}_1\theta}$ or greater than $w_{{}_1\theta}$, respectively.

Definition of set σ. Let λ be a fixed positive integer such that $r \leqq \lambda \leqq n$,

where r and n are as defined above for (1.2). We shall call $x_{1\cdots\lambda}$ a set σ [relative to the Kellogg solution w of (1.2)], if and only if $\varphi_p(1/x) \leqq \varphi_p(1/w)$ for every positive integer p such that $r \leqq p \leqq \lambda$.

Proof of (4.3). In I we showed that if $X_{1\cdots p}$, $r \leqq p \leqq n - 1$, is any set of p positive integers such that

$$\Sigma_{p,r}(1/X) < b/a, \qquad a \equiv [(c+1)b - 1],$$

then[13]

$$\Sigma_{p,r}(1/X) \leqq \Sigma_{p,r}(1/w), \tag{4.6}$$

where w is the Kellogg solution of (1) of I. Hence (4.6) holds in particular when X is any E-solution of (1.2) and w is the Kellogg solution of (1.2), the first $n - 1$ elements w_i of the two w's in question being identical [cf. (26) of I and (4) here]. Relation (4.3) is the case $p = n - 1$ of (4.6).

Remark. Since (4.6) (a relation from I) happens to hold here, it is not necessary for us to make an induction here parallel to the main induction of I. This explains why we use n rather than an induction integer k, $r \leqq k \leqq n - 1$, as the subscript of φ in (4.5).

The completion of the proof of the remarkable properties for w will not be difficult after we establish the following lemma, which resembles certain parts of Lemma 9 of I.

LEMMA 4. *Let f $(\equiv f_{1\cdots n})$ stand for our E-solution X $(\neq w)$ or any intermediate set of X under transformation* (4.5). *Then application of* (4.5) *to f is such that* (3.1) *and the following relations hold:*

$$\Sigma_{p,r}(1/f') \leqq \Sigma_{p,r}(1/w) \qquad (r > 1;\ p = r, \cdots, {}_1\theta - 1), \tag{4.7}$$

$$\Sigma_{p,r}(1/f') \leqq \Sigma_{p,r}(1/f) \qquad (r > 1;\ p = {}_1\theta, \cdots, n - 1). \tag{4.8}$$

If we prove the case $f = X$ of Lemma 4, the other cases will follow as did the corresponding further cases in I (cf. Lemma 9).

Proof that (3.1) *holds when $f = X$.* Using all cases $p = r, \cdots, n - 1$ of (4.6), we observe that $f = X$ is a set σ, so that $\theta_1 = q_1 < {}_1\theta = {}_1q$. Since X is an E-solution, $f_{\theta_1} \leqq f_{{}_1\theta}$. Consequently, (3.1) holds (cf. Lemma 1a of I).

Proof that (4.7) *holds when $f = X$.* By (4.5), $f'_t = f_t \geqq w_t$ for all positive integral values of t from 1 to ${}_1\theta - 1$ except $t = \theta_1$, and $w_{\theta_1} \leqq f'_{\theta_1} < f_{\theta_1}$. Hence (4.7) holds when $f = X$.

Proof that (4.8) *holds when $f = X$.* This inequality is the present analog (cf. footnote 10) of (37) of I. Since (37) and (46) of I are equivalent, in order to

[13] It is to be noted that the hypothesis in I that X is an E-solution of equation (1) of I was not used (in full) in proving (4.6). This hypothesis was used only when the value n was assigned to p.

prove the case $f = X$ of (4.8) it suffices to establish the analog here of (46) of I. This analog is readily found to be[14]

$$\Sigma'_{n-2,r-1}\Sigma'_{p-2,r-2} \geqq (\Sigma'_{n-2,n-2} + \Sigma'_{n-2,r-2})\Sigma'_{p-2,r-1}, \tag{4.9}$$

where[15] $p = {}_1q, \cdots, n-1$. Our method of proving (4.9) is *to show first that this relation is true for* $p = n - 1$, *and then to prove that if* (4.9) *is true for* $p = k$, *where* k *is any admissible value of* p *except its smallest value*, ${}_1q$, *then* (4.9) *is also valid for* $p = k - 1$. We have not been able to prove directly the inequality which one obtains by putting $p = n - 1$ in (4.9). However, by reducing the terms of each $\Sigma'_{\alpha,\beta}$ of this relation to a common denominator and then cancelling the denominators, we obtain the following equivalent of the case $p = n - 1$ of (4.9), which we shall establish:

$$\Sigma'_{n-2,\,n-r-1}(X)\Sigma'_{n-3,\,n-r-1}(X) \geqq [\Sigma'_{n-2,\,n-r}(X) + 1]\,\Sigma'_{n-3,\,n-r-2}(X). \tag{4.10}$$

Employing the identities[16]

$$\Sigma'_{n-2,t}(X) \equiv \Sigma'_{n-3,t}(X) + X_n\Sigma'_{n-3,t-1}(X) \qquad (t = n - r - 1, n - r),$$

we find that (4.10) reduces to

$$[\Sigma'_{n-3,\,n-r-1}(X)]^2 \geqq [\Sigma'_{n-3,\,n-r}(X) + 1]\Sigma'_{n-3,\,n-r-2}(X). \tag{4.11}$$

The symbols in (4.11) are all positive except for $r = 2$; in this case $\Sigma'_{n-3,\,n-r}(X) = 0$ (cf. the definition of $\Sigma_{i,j}(x)$ in §1 and of $\Sigma'_{i,j}(x)$ in footnote 14) and (4.11) is equivalent to

$$X_{i_1}^2 \cdots X_{i_{n-3}}^2 \geqq \Sigma'_{n-3,\,n-4}(X), \tag{4.12}$$

where $X_{i_1}, \cdots, X_{i_{n-3}}$ are the $n - 3$ elements of X which are under consideration, and $i_1 < i_2 < \cdots < i_{n-3}$. Since X is by hypothesis an E-solution of (1.2), the right member of (4.12) does not exceed $(n-3)X_{i_1} \cdots X_{i_{n-3}}$. From this fact and the inequalities $X_1 \geqq 1, X_t \geqq 2, t = 2, \cdots, n$ (which are obviously true when $r = 2$) one readily finds that in order to establish (4.12) it suffices to prove that

$$2^{n-4} \geqq n - 3 \qquad (n \geqq 4). \tag{4.13}$$

Here the equality sign holds for $n = 4, 5$, and the inequality sign holds for $n > 5$. Hence (4.13), the case $r = 2$ of (4.10), is true.

[14] Here, as in the sequel, $\Sigma'_{\alpha,\beta} \equiv \Sigma'_{\alpha,\beta}(1/X)$ is the β-th elementary symmetric function of all of the reciprocals $1/X_1, \cdots, 1/X_{\alpha+2}$ except $1/X_{q_1}$ and $1/X_{{}_1q}$.

[15] Since $w_i = 1$ $(i = 1, \cdots, r-1)$, ${}_1q \geqq r$. If ${}_1q = n$, X_n is the only element of class B in X, and every transformation that we make in passing from X to w will accord with (3.1), so that we shall not need to consider (4.9). In other words, no E-solution X in which $X_{1\ldots(n-1)}$ is non-transformable fails to accord with the theorem that w has the remarkable properties. In the present proof, then, we suppose $X_{1\ldots(n-1)}$ transformable so that ${}_1q \leqq n - 1$.

[16] By the hypothesis that ${}_1q \leqq n - 1$, every $\Sigma'_{n-2,\beta}(X)$ under consideration contains X_n.

When $r > 2$ *in* (4.10), we let $(u^s) \equiv \Sigma'_{n-3,s}(X^u)$. Hence to prove (4.11), it suffices to show that, with $s \equiv n - r - 1 \geqq 1$ [cf. the hypothesis $n \geqq r + 2$ in (1.2)]

$$(1^s)^2 \geqq (1^{s+1})(1^{s-1}) + (1^{s-1}). \tag{4.14}$$

By use of (47) and (48) of I, with $s_1 = s_2 = n - r - 1 = s \geqq 1$, we find that

$$(1^s)^2 - (1^{s+1})(1^{s-1}) = [(2^s) + (2^{s-1}1^2) + \cdots]. \tag{4.15}$$

Under present hypotheses (j^t), $j = 1, 2$, is a function of $n - 3 \geqq s + 1$ variables X_i. Hence (2^s) and $(2^{s-1}1^2)$ are both positive and

$$(2^{s-1}1^2) \geqq (1^{s-1}),$$

while no omitted term in the right member of (4.15) is negative. Consequently (4.14) holds in the sense of $>$.

Final step of the proof of (4.9). We have proved (4.9) for $p = n - 1$. If the set of numbers ${}_1q, \cdots, n - 1$ contains only one number, our argument is complete; otherwise, we assume that (4.9) is true for $p = k > {}_1q$ (and $< n$; cf. footnote 14) so that

$$\Sigma'_{k-2,r-2}[\Sigma'_{k-2,r-1}]^{-1} \geqq [\Sigma'_{n-2,r-2} + \Sigma'_{n-2,n-2}]\,[\Sigma'_{n-2,r-1}]^{-1} \qquad (r > 1). \tag{4.16}$$

With (4.16) holding, we wish to prove that if R is the right member of (4.16), then (in our induction from k to $k - 1$)

$$\Sigma'_{k-3,r-2}[\Sigma'_{k-3,r-1}]^{-1} \geqq R. \tag{4.17}$$

To establish (4.17), it suffices to show that the left member of this relation is not less than that of (4.16), or that

$$\Sigma'_{k-2,r-1}\Sigma'_{k-3,r-2} \geqq \Sigma'_{k-2,r-2}\Sigma'_{k-3,r-1} \qquad (k > {}_1q \geqq r). \tag{4.18}$$

Since in the case ${}_1q = r$, $(X'_1 \cdots X'_r)^{-1} < (X_1 \cdots X_r)^{-1} \leqq (w_1 \cdots w_r)^{-1}$ here as in I [cf. equation (39) of I], we may, and do, assume that if ${}_1q = r$, then $k \geqq r + 2$ and that if ${}_1q > r$, then $k \geqq {}_1q + 1$. Therefore k is never less than $r + 2$, and each member of (4.18) is positive. To establish (4.18) we use in it the identities

$$\Sigma'_{k-2,t} = \Sigma'_{k-3,t} + X_n^{-1}\Sigma'_{k-3,t-1} \qquad (t = r - 1, r - 2),$$

and obtain the following equivalent of (4.18)

$$(\Sigma'_{k-3,r-2})^2 \geqq \Sigma'_{k-3,r-3}\Sigma'_{k-3,r-1} \qquad (k \geqq r + 2);$$

or, in our partition notation, with $\Sigma'_{k-3,u} \equiv (1^u)$,

$$(1^s)^2 \geqq (1^{s-1})(1^{s+1}), \quad s = (r - 2) \geqq 0, \quad (k - 3) \geqq (s + 1). \tag{4.19}$$

Relation (4.19) is obviously true when $s = 0$, and it is true by (4.14) when $s > 0$. Hence (4.9), the case $f = X$ of (4.8), is true.

If $X'_{1\,\ldots\,(n-1)}$ contains at least one element of each of the classes A', B',

application of (4.5) with $f = X'$ will yield this case of (4.8) [cf. the proof of the case $f = X'$ of (56) of I], etc., until we arrive at a set $X^{(t)}$ such that $X^{(t)}_{1\cdots(n-1)}$ does not contain an element of both the classes $A^{(t)}$, $B^{(t)}$, and $X^{(t-1)}_{1\cdots(n-1)}$ contains an element of each of the classes $A^{(t-1)}$, $B^{(t-1)}$, while (4.8) holds for $f = X$, X', $\cdots$, $X^{(t-1)}$ and, therefore, for $f' = X^{(t)}$. Hence (4.8) is true.

Completion of the proof of the remarkable properties for w. Relations (4.7), (4.8), and (4.3) imply that every set f which we consider is a set σ. Consequently, from the last paragraph above,

$$X^{(t)}_i \geqq w_i \qquad (i = 1, \cdots, n-1) . \tag{4.20}$$

We desire to prove that the sign $>$ holds in (4.20) for at least one of the values of i. Suppose that the equality sign holds in (4.20) for every admissible value of i. Then, since $X^{(t)}$ satisfies the case $r > 1$ of (1.2), $X^{(t)} = w$. We shall reach a contradiction. Here, as in (29) of I,

$$\Sigma_{p,r}(1/X) \leqq (bX_1 \cdots X_p - 1)/(aX_1 \cdots X_p) \qquad (p = {}_1q, \cdots, n-1) . \tag{4.21}$$

With $p = n - 1$, relations (4.21), (4.8) and our hypothesis that $X^{(t)} = w$ now imply that

$$\Sigma_{n-1,r}(1/w) \leqq \Sigma_{n-1,r}(1/X) . \tag{4.22}$$

Consequently it follows from (4.21), (4.22), and the fact that (28) of I holds here that

$$w_1 \cdots w_{n-1} \leqq X_1 \cdots X_{n-1} . \tag{4.23}$$

However, every transformation that we made in passing from X (in which $X_{1\cdots(n-1)}$ was supposed to be transformable) to $X^{(t)} = w$ affected exactly two of the first $n-1$ elements of one of the sets $X, X', \cdots, X^{(t-1)}$ and increased the product of these two, so that $w_1 \cdots w_{n-1} > X_1 \cdots X_{n-1}$, which contradicts (4.23). Hence the sign $>$ holds in (4.20) for at least one of the specified values of i.

That w has the remarkable properties now follows readily. Since $X^{(t)}$ satisfies the case $r > 1$ of (1.2), and (4.20) holds in the sense just described,

$$X^{(t)}_n < w_n . \tag{4.24}$$

By the definition of t, $X_n = X^{(t)}_n$. Therefore $X_n < w_n$ and w *has the first remarkable property*. Further, from (4.20) and (4.24) it follows (cf. footnote 15) that all applications of (4.5) which our procedure employs in passing from $X^{(t)}$ to w accord with (3.1). Hence from Lemma 3 of I it follows that w *has the second remarkable property*.

Proof of (4.4). The first remarkable property for w [cf. the first sentence below (4.2)] has just been established without proving that (4.4) is valid.[17]

[17] Most of §20 of I could be replaced by argument of the type that is used in the two paragraphs just before the proof of (4.4).

That (4.4) is true may be proved as follows. In every set that we consider $X_{1 \ldots (n-1)}$ is transformable, and by (4.5) and (4.8) every transformation that we use in passing from X to $X^{(i)}$ is such that

$$\varphi_n(1/f') = \varphi_n(1/f)\,, \qquad \Sigma_{n-1,r}(1/f') \leqq \Sigma_{n-1,r}(1/f)\,,$$

the first f admitted here being X. Hence from (4.2) it follows that in each such transformation

$$\psi(1/f') \geqq \psi(1/f)\,. \tag{4.25}$$

Then every further application of (4.5) that is needed to carry $X^{(i)}$ (which was shown to be different from w in the second paragraph above) into w is such that

$$f'_{\theta_1} < f_{\theta_1}\,, \qquad f'_p = f_p \qquad (p = 1, \cdots, \theta_1 - 1, \theta_1 + 1, \cdots, n-1)\,,$$

so that here (4.25) holds in the sense of $>$. Hence $\psi(1/X) < \psi(1/w)$.

The statements of the results of Part 1 are obtained by setting $\lambda_t = 0$ $(t = r+1, \cdots, s)$ in Theorem 9.

Part 2. The maximum number and the class of maximum numbers that we associate with equation (1.3)

5. **Statement of results analogous to those in I.** We consider equation (1.3) as conditioned by (1.4). The Kellogg solution w of this equation is $x = w$, where [cf. (23) of I]

$$\begin{aligned}
&w_p = 1\ (p = 1, \cdots, r-1)\,, \qquad w_r = c + 1\,,\\
&w_{p+1} = a[\Sigma_{p,\,p-r+1}(w) + \lambda_{r+1}\Sigma_{p,\,p-r}(w) + \lambda_{r+2}\Sigma_{p,\,p-r-1}(w)\\
&\qquad\qquad + \cdots + \lambda_s\Sigma_{p,\,p-s+1}(w)] + 1 \qquad (p = r, \cdots, n-2)\,,\\
&w_n = a[\Sigma_{n-1,\,n-r}(w) + \lambda_{r+1}\Sigma_{n-1,\,n-r-1}(w) + \lambda_{r+2}\Sigma_{n-1,\,n-r-2}(w)\\
&\qquad\qquad + \cdots + \lambda_s\Sigma_{n-1,\,n-s}(w)]\,.
\end{aligned} \tag{5}$$

THEOREM 5. *The solution w just defined has the remarkable properties relative to the equation* (1.3) *that we consider* [*including the case* $\lambda_t = 0$ $(t = r+1, \cdots, s)$ *of* I].

6. **Method of procedure.** In proving Theorem 5, we use the methods of I, and therefore of Part 2 of II. We define the left member of (1.3) to be $\varphi_n(1/x)$ and write

$$\begin{aligned}
\varphi_p(1/x) \equiv\ &\Sigma_{p,r}(1/x) + \lambda_{r+1}\Sigma_{p,r+1}(1/x) + \lambda_{r+2}\Sigma_{p,r+2}(1/x)\\
&+ \cdots + \lambda_s\Sigma_{p,s}(1/x) \qquad (p = r, \cdots, n)\,.
\end{aligned} \tag{6}$$

Then an equivalent of (1.3) is

$$\varphi_{n-1}(1/x) + (1/x_n)\psi(1/x) = b/a\,, \tag{6.1}$$

where

$$\begin{aligned}\psi(1/x) \equiv \Sigma_{n-1,\,r-1}(1/x) + \lambda_{r+1}\Sigma_{n-1,\,r}(1/x) + \lambda_{r+2}\Sigma_{n-1,\,r+1}(1/x)\\ + \cdots + \lambda_s\Sigma_{n-1,\,s-1}(1/x)\,.\end{aligned} \tag{6.2}$$

To establish the first remarkable property for w, it suffices to prove that if X is any E-solution of our equation (1.3) except its Kellogg solution, w, then[18] [cf. (4.3) and (4.4)]

$$\varphi_{n-1}(1/X) \leqq \varphi_{n-1}(1/w)\,, \tag{6.3}$$

$$\psi(1/X) < \psi(1/w)\,. \tag{6.4}$$

Further, if in proving these relations, we apply (as our method directs us to do) one or more transformations of the type defined in (6.5) below which satisfy (3.1), and only such transformations, we shall establish the second remarkable property for w (cf. Lemma 3 of I).

To avoid redundancy and yet point the way to the conclusion of the remarkable properties for w, we next present a rough parallel (somewhat amplified in places) of Part 1 of II. In this parallel we omit Lemmas 7 and 8 of II, which need not be changed here.

The analogs here of (26a), (30a), (32a) of II are (5), (6.3), (6.4), respectively; and the present analogs of (28a), (29a) of II are these respective relations themselves except that now $\varphi_p(1/x)$ is defined by (6).

Sets σ, τ. In the definition of set σ in §4, merely replace (1.2) and the $\varphi_p(1/x)$ of (4.1) by (1.3) and the $\varphi_p(1/x)$ of (6), respectively, to obtain the present definition of set σ. Set τ is defined here as it was on page 890 of I.

The new transformation. Using the notation of Lemma 4 above, and also of Lemma 9 of I, we suppose that[19] $f_{1\ldots k}$, $r < k \leqq n$, where $f \equiv X^{(\alpha)}$, contains at least one element of each of the classes $A^{(\alpha)}$, $B^{(\alpha)}$. Then our transformation from $f_{1\ldots k}$ to $f'_{1\ldots k}$ is t_5 or t_6,

$$\begin{aligned}(t_5)\; f'_p = f_p\ (p \neq \theta_1, {}_1\theta, 1 \leqq p \leqq k)\,, \quad f'_{\theta_1} = w_{\theta_1}\,, \quad \varphi_k(1/f') = \varphi_k(1/f)\,,\\ (t_6)\; f'_p = f_p\ (p \neq \theta_1, {}_1\theta, 1 \leqq p \leqq k)\,, \quad f'_{{}_1\theta} = w_{{}_1\theta}\,, \quad \varphi_k(1/f') = \varphi_k(1/f)\,,\end{aligned} \tag{6.5}$$

according as t_5 defines $f'_{{}_1\theta}$ to be not greater than $w_{{}_1\theta}$ or greater than $w_{{}_1\theta}$, respectively.

[18] A proof which in our procedure is essentially equivalent to, and slightly shorter than, our proof of (6.3) and (6.4) consists in showing that in the notation of §4, (4.20) and (4.24) hold here. In §4 we gave the analogs of both of the proofs in question. Were we to give here all details of our treatment of (1.3), we would not feel content to conclude our argument without observing its obvious implication that (6.4) is true.

[19] Just as we used the relations $r < k \leqq n$ in Lemma 9 of I, we may employ them here.

Lemma 6. (*Analog of Lemma* 9 *of* I).[20] (i) *If* $X_{1\ldots k}$ *is a set* τ, $X_{1\ldots k}$ *is transformable.*[21] (ii) *If* $f_{1\ldots k}$ *is a set* τ *or a transformable set* σ *for which* $r < k \leqq n$ *and if* t *is a positive integer, application of* (6.5) *to* $f_{1\ldots k}$ *yields a set* $f'_{1\ldots k}$ *such that* (3.1) *holds, and*

$$\varphi_p(1/f') \leqq \varphi_p(1/w) \qquad (p = r, \cdots, {}_1\theta - 1)\,, \tag{6.6}$$

$$\varphi_p(1/f') \leqq \varphi_p(1/f) \qquad (p = {}_1\theta, \cdots, k - 1)\,, \tag{6.7}$$

$$\varphi_k(1/f') = \varphi_k(1/f)\,. \tag{6.8}$$

Partial proof of Lemma 6. The proof of (i) is the same as that of the corresponding case in Lemma 6 of I. Concerning (ii), the case $f = X$ of (3.1) follows from the relation $X_{q_1} \leqq X_{{}_1q}$, which is obviously true, and Lemma 1a of I; the case $f = X$ of (6.6) follows from the facts that ${}_1q \geqq r$ and $X_i \geqq w_i$ for $i = 1, \cdots, {}_1q - 1$ [cf. the proof of (35) of I]; (6.8) is true by (6.5).

In §7 we establish the case $f = X$ of (6.7). Then the rest of the proof of Lemma 6 can be made by argument of a type that was used on pages 899 and 900 of I.

7. **The analog of** (46) **of I. How to complete the proof of Theorem** 5. From the case $f = X$ of (6.7) and (6.8), one can derive (7) below just as we obtained (46) of I from the relations (37) and (38) of I:

$$\begin{aligned} &[\Sigma'_{k-2,r-1} + \lambda_{r+1}\Sigma'_{k-2,r} + \lambda_{r+2}\Sigma'_{k-2,r+1} + \cdots + \lambda_s\Sigma'_{k-2,s-1}] \\ &\qquad [\Sigma'_{p-2,r-2} + \lambda_{r+1}\Sigma'_{p-2,r-1} + \lambda_{r+2}\Sigma'_{p-2,r} + \cdots + \lambda_s\Sigma'_{p-2,s-2}] \\ &\geqq [\Sigma'_{k-2,r-2} + \lambda_{r+1}\Sigma'_{k-2,r-1} + \lambda_{r+2}\Sigma'_{k-2,r} + \cdots + \lambda_s\Sigma'_{k-2,s-2}] \\ &\qquad [\Sigma'_{p-2,r-1} + \lambda_{r+1}\Sigma'_{p-2,r} + \lambda_{r+2}\Sigma'_{p-2,r+1} + \cdots + \lambda_s\Sigma'_{p-2,s-1}] \\ &\qquad\qquad (k > p \geqq {}_1q \geqq r,\ s > r). \end{aligned} \tag{7}$$

We shall presently show that (7) is equivalent to[22]

$$\Sigma(\lambda_{i+1}\lambda_{j+2} - \lambda_{i+2}\lambda_{j+1})[\Sigma'_{k-2,i}\Sigma'_{p-2,j} - \Sigma'_{k-2,j}\Sigma'_{p-2,i}] \geqq 0, \tag{7.1}$$

where the summation extends over $i = r - 1, \cdots, s - 1; j = r - 2, \cdots, i - 1$ and $\lambda_{r-1} = \lambda_{s+1} = 0$, $\lambda_r = 1$. This equivalence may be proved as follows. Certainly (7) is equivalent to

$$\Sigma\lambda_{i+1}\lambda_{j+2}(\Sigma'_{k-2,i}\Sigma'_{p-2,j} - \Sigma'_{k-2,j}\Sigma'_{p-2,i}) \geqq 0,$$

[20] Lemma 6 and transformation (6.5) include the lemma and transformation, respectively, which would be used in a precise parallel here of Part 1 of II.

[21] In Lemma 6, $X_{1\ldots k}$ must be transformable for at least one value of k such that $r < k < n$; for, as was stated in footnote 15, no E-solution X in which $X_{1\ldots(n-1)}$ is nontransformable needs to be considered.

[22] The sign = would hold here if $n = 4$, $p = 2$, $k = 3$, $q_1 = 1$, ${}_1q = 2$, and $\varphi_4(1/X) = \Sigma_{4,1}(1/X) + \Sigma_{4,2}(1/X) + \Sigma_{4,3}(1/X)$. This example shows that the sign $<$ does not always hold in (37a) and (46a) of II. In both of these relations $<$ should be replaced by $\leqq$.

where the summation extends over all $i, j = r - 2, \cdots, s - 1$. Further, the sum of all terms that one obtains here by setting $i = j$ is zero. Consequently, consideration of the two cases $i > j$ and $i < j$ now leads to the conclusion that (7.1) is equivalent to (7).

Proof of (7.1). On account of (1.4), we only need to prove that

$$\Sigma'_{k-2,i}\Sigma'_{p-2,j} - \Sigma'_{k-2,j}\Sigma'_{p-2,i} \geqq 0. \tag{7.2}$$

There are two cases to be considered: (1) where Lemma 8a of II applies to (7.2) with $u = k - 2$, $v = p - 2$, $\gamma = i$, $t = j$; (2) where this lemma does not apply to (7.2).

(1) Here (7.2) holds in the sense of $>$, and (7.1) is true.

(2) Since $k > p$ and $i > j$ by hypothesis, we conveniently subdivide this case into two other cases, as follows: (α) when $j < 0$, as it is if $j = (r - 2)$ and $r = 1$; (β) when $i > (p - 2)$. In case (α) [(β)] each term [the last term] in the left member of (7.2) is zero. Therefore (7.2) holds in case (2).

How to complete the proof of theorem 5. With Lemma 6 holding (cf. the last sentence of §6), we can make the induction for (6.3) just as the similar induction for (30) of I was made in I (cf. in particular the beginning of §14 and the last two paragraphs of both §16 and §18, of I). Then argument of §4 leads to the analogs here of (4.20) and (4.24) and thus establishes the first remarkable property for w. Next we observe that throughout the proof of Lemma 6 the relation [cf. (58) of I]

$$f'_{\theta_1} \leqq f_{1\theta}$$

is used, so that each transformation that we employ in this proof accords with (3.1). Consequently, the second remarkable property for w follows.

8. **Our most general result relative to equations of the form** (1.3). The methods by which one proves Theorems 2a, 3a, 4a, 5a of II suffice to prove the following

THEOREM 8. *Let $\varphi_p(1/x)$ be as it is in* (6). *Suppose that a, b, and μ, where $r \leqq \mu \leqq n - 1$, are given positive integers, with a and b, $b \leqq a$, relatively prime, and that there exists a set of n numbers $w \equiv (w_1, \cdots, w_n)$ with the following properties:*

(1) *it is an E-solution of the equation*

$$\varphi_n(1/x) = b/a, \quad 1 \leqq r < s \leqq n, \quad \lambda\text{'s as in } (1.4);$$

$$\varphi_p(1/w) = (bw_1 \cdots w_p - 1)/(aw_1 \cdots w_p) \tag{2}$$

for every positive integral value of p for which $\mu \leqq p \leqq (n - 1)$ and for no positive integral value of p less than μ;

(3) *if $x = X$, where $X \neq w$, is an E-solution of the equation in* (1), *then for every positive integral value of p such that $r \leqq p \leqq \mu$ for which $X_{1\cdots p} \neq w_{1\cdots p}$ the relation $\varphi_p(1/x) \leqq \varphi_p(1/w)$ holds.*

Then w is the Kellogg solution of the equation in (1) and w has the remarkable properties relative to this equation.

The following corollary shows that Theorem 8 has content for a case in which $\mu > r$.

COROLLARY 8. *The Kellogg solution of the equation*

$$\Sigma_{n,1}(1/x) + 2\,\Sigma_{n,2}(1/x) + 3\Sigma_{n,3}(1/x) = 5/11 \qquad (n > 4)$$

is $x = w$, *where* $w_1 = 3$, $w_2 = 14$, *and*

$$w_{p+1} = 11[\Sigma_{p,p}(w) + 2\Sigma_{p,p-1}(w) + 3\Sigma_{p,p-2}(w)] + 1 \quad (p = 2, \cdots, n-2),$$

$$w_n = 11[\Sigma_{n-1,n-1}(w) + 2\Sigma_{n-1,n-2}(w) + 3\Sigma_{n-1,n-3}(w)].$$

Here the μ *of Theorem* 8 *has the value* $2 > r = 1$.

Part 3. The maximum number and the class of maximum numbers that we associate with equation (1.5)

9. Statement of results analogous to those in Part 1. We consider equation (1.5) as conditioned by (1.4). The Kellogg solution w of this equation is

$$w_i \ (i = 1, \cdots, n-1) \text{ as in (5)},$$

$$w_n = a[\Sigma_{n-1,n-r}(w) + \lambda_{r+1}\Sigma_{n-1,n-r-1}(w) + \lambda_{r+2}\Sigma_{n-1,n-r-2}(w) + \cdots + \lambda_s\Sigma_{n-1,n-s}(w) + 1], \tag{9}$$

where the λ's satisfy (1.4). We give below the main step in the proof of the case $\lambda_t > 0$ $(t = r+1, \cdots, s)$ of the following

THEOREM 9. *The solution* w *in* (9) *has the remarkable properties relative to its associated equation*[23] (1.5).

10. Proof of the case $\lambda_t > 0$ of Theorem 9. We use the methods of Part 1, §4. We define the left member of our equation (1.5) to be the new function $\varphi_n(1/x)$, and write

$$\varphi_p(1/x) \equiv \Sigma_{p,r}(1/x) + \lambda_{r+1}\Sigma_{p,r+1}(1/x) + \lambda_{r+2}\Sigma_{p,r+2}(1/x) + \cdots + \lambda_s\Sigma_{p,s}(1/x) + \Sigma_{p,n}(1/x) \qquad (p = r, \cdots, n), \tag{10}$$

so that the present $\varphi_p(1/x)$ is formally the same as that of (6) except for $p = n$. An equivalent of (1.5) is

$$\varphi_{n-1}(1/x) + (1/x_n)\psi(1/x) = b/a, \tag{10.1}$$

where

$$\psi(1/x) \equiv \Sigma_{n-1,r-1}(1/x) + \lambda_{r+1}\Sigma_{n-1,r}(1/x) + \lambda_{r+2}\Sigma_{n-1,r+1}(1/x) + \cdots + \lambda_s\Sigma_{n-1,s-1}(1/x) + \Sigma_{n-1,n-1}(1/x).$$

[23] The case $\lambda_t = 0$ of Theorem 9 was disposed of in Part 1. The only case that we need to consider here is that in which $\lambda_t > 0$ (cf. footnote 7).

Now we assume X to be any E-solution of our equation (1.5) except its Kellogg solution, w. In order to establish the first remarkable property for w here, it suffices to prove that under our present hypotheses the inequalities [cf. (4.3), (4.4)]

$$\varphi_{n-1}(1/X) \leqq \varphi_{n-1}(1/w), \tag{10.2}$$

$$\psi(1/X) < \psi(1/w) \tag{10.3}$$

hold, so that $X_n < w_n$ [cf. (10.1)]. If in establishing (10.2) and (10.3) we transform X into w by one or more transformations, everyone of which accords with (3.1), we shall prove for w the second remarkable property, and thus obtain the case $\lambda_t > 0$ of Theorem 9. After we define the *new transformation* and *set* σ, we shall give (in §11) the proof of the analog here of (4.9). All further details in the proof of the case $\lambda_t > 0$ of Theorem 9 will then be obvious from the first three paragraphs below (4.19).

Definition of the transformation. This is formally the same as the definition of the transformation of §4, the symbols φ_n, w, X, f here relating to the case $\lambda_t > 0$ of (1.5) rather than (1.2).

Definition of set σ. This is also formally the same as the definition of set σ in §4 except that here the present case of (1.5) and the φ_p of (10) replace (1.2) and the φ_p of (4.1), respectively, there.

11. **Analog of** (4.9). This relation is readily found to be[23]

$$\begin{aligned} &[\Sigma'_{n-2,r-1} + \lambda_{r+1}\Sigma'_{n-2,r} + \lambda_{r+2}\Sigma'_{n-2,r+1} + \cdots + \lambda_s\Sigma'_{n-2,s-1}] \\ &\qquad [\Sigma'_{p-2,r-2} + \lambda_{r+1}\Sigma'_{p-2,r-1} + \lambda_{r+2}\Sigma'_{p-2,r} + \cdots + \lambda_s\Sigma'_{p-2,s-2}] \\ &> [\Sigma'_{n-2,r-2} + \lambda_{r+1}\Sigma'_{n-2,r-1} + \lambda_{r+2}\Sigma'_{n-2,r} + \cdots + \lambda_s\Sigma'_{n-2,s-2} + \Sigma'_{n-2,n-2}] \\ &\qquad [\Sigma'_{p-2,r-1} + \lambda_{r+1}\Sigma'_{p-2,r} + \lambda_{r+2}\Sigma'_{p-2,r+1} + \cdots + \lambda_s\Sigma'_{p-2,s-1}] \\ &\qquad\qquad (p = {}_1q, \cdots, n-1;\ \lambda_t > 0). \end{aligned} \tag{11}$$

It can be proved[24] that (11) is equivalent to (11.1) in the same way that we showed (7) to be equivalent to (7.1):

$$\begin{aligned} &\Sigma(\lambda_{i+1}\lambda_{j+2} - \lambda_{i+2}\lambda_{j+1})(\Sigma'_{n-2,i}\Sigma'_{p-2,j} - \Sigma'_{n-2,j}\Sigma'_{p-2,i}) \\ &\quad > \Sigma'_{n-2,n-2}[\Sigma'_{p-2,r-1} + \lambda_{r+1}\Sigma'_{p-2,r} + \lambda_{r+2}\Sigma'_{p-2,r+1} + \cdots + \lambda_s\Sigma'_{p-2,s-1}], \end{aligned} \tag{11.1}$$

where the summation is to be extended over $i = r-1, \cdots, s-1$; $j = r-2, \cdots, i-1$, and $\lambda_{r-1} = \lambda_{s+1} = 0$, $\lambda_r = 1$. By (1.4) the first parenthesis under the summation sign is not negative for any admissible pair of values of i and j. We prove that (11.1) is true in two steps. First we show that (11.1) holds for

[24] It is interesting to note that if $r = 1$ and $\lambda_t = 0$ ($t = r+1, \cdots, s$), (11) does not hold. This is the case which we gave special treatment in Part 1. It is also to be noted that the sign $>$, rather than $\geqq$, is used in (11).

$p = n - 1$. Then we prove that if (11) holds for $p = k$, where k is any admissible value of p except its smallest[25] value ${}_1q$, (11) is true also for $p = k - 1$.

First step. We treated the case $s = r = 1$ of (1.5) in §§2, 3; the case $s = r > 1$, in §4. Under our present hypothesis that $\lambda_t > 0$ $(t = r + 1, \cdots, s)$ we need now to dispose of the following cases of (1.5): (i) $s > r = 1$; (ii) $s > r > 1$. This we now proceed to do.

(i) The case $s > r = 1$. If $s = 2, 3, m$, $(3 < m \leqq n - 2)$, and $p = n - 1$, (11.1) reduces to (11.2), (11.3), (11.4), respectively:

$$\lambda_2^2(\Sigma'_{n-2,1} - \Sigma'_{n-3,1}) > \Sigma'_{n-2,n-2}(1 + \lambda_2\Sigma'_{n-3,1}), \tag{11.2}$$

$$\begin{aligned}[t] &[(\lambda_2^2 - \lambda_3)(\Sigma'_{n-2,1} - \Sigma'_{n-3,1}) + \lambda_2\lambda_3(\Sigma'_{n-2,2} - \Sigma'_{n-3,2}) \\ &+ \lambda_3^2(\Sigma'_{n-2,2}\Sigma'_{n-3,1} - \Sigma'_{n-2,1}\Sigma'_{n-3,2})] > \Sigma'_{n-2,n-2}(1 + \lambda_2\Sigma'_{n-3,1} + \lambda_3\Sigma'_{n-3,2}),\end{aligned} \tag{11.3}$$

$$\begin{aligned} &[A + \lambda_m\lambda_2(\Sigma'_{n-2,m-1} - \Sigma'_{n-3,m-1}) + \lambda_m\lambda_3(\Sigma'_{n-2,m-1}\Sigma'_{n-3,1} \\ &- \Sigma'_{n-2,1}\Sigma'_{n-3,m-1}) + \cdots + \lambda_m^2(\Sigma'_{n-2,m-1}\Sigma'_{n-3,m-2} - \Sigma'_{n-2,m-2}\Sigma'_{n-3,m-1})] \\ &> \Sigma'_{n-2,n-2}(1 + \lambda_2\Sigma'_{n-3,1} + \lambda_3\Sigma'_{n-3,2} + \cdots + \lambda_m\Sigma'_{n-3,m-1}),\end{aligned} \tag{11.4}$$

where A is the sum of those terms of the left member of (11.1) that are obtained by assigning to i values $> r - 2$ and $< m - 1 = (s - 1)$.

Proof of (11.2). The left member of (11.2) equals $\lambda_2^2X_n^{-1}$ and $\Sigma'_{n-2,n-2} = (PX_n)^{-1}$, where $P \equiv X_{i_1} \cdots X_{i_{n-2}}$ [cf. the notation of (4.12)]. After reducing all terms of (11.2) to a common denominator, we find that this inequality is equivalent to

$$P^2\lambda_2^2 > P + \lambda_2\Sigma'_{n-3,n-4}(X). \tag{11.5}$$

If $n = 4$, the smallest admissible value of n when $s = 2$, (11.5) reduces to

$$X_{i_1}^2\lambda_2^2 > X_{i_1} + \lambda_2.$$

Since $r = 1$, every element X_t of X exceeds unity, and $\lambda_2 \geqq 1$; hence the last displayed relation, and therefore the case $n = 4$ of (11.5), is true. Suppose $n > 4$. Since λ_2 is a positive integer and $X_1 \leqq X_2 \leqq \cdots \leqq X_n$, we shall establish (11.5) if we prove that $P > [1 + (n - 3)/X_{i_1}]$ or, indeed, that $2^{n-2} > n - 1$. This inequality is true for $n > 3$ and therefore for the cases under consideration. Consequently (11.5) and its equivalent (11.2) hold.

Proof of (11.3). We shall obtain the desired result by establishing the inequalities

$$\Sigma'_{n-2,2} - \Sigma'_{n-3,2} > \Sigma'_{n-2,n-2}(1 + \Sigma'_{n-3,1}), \tag{11.6}$$

$$\Sigma'_{n-2,2}\Sigma'_{n-3,1} - \Sigma'_{n-2,1}\Sigma'_{n-3,2} > \Sigma'_{n-2,n-2}\Sigma'_{n-3,2}. \tag{11.7}$$

[25] As in the corresponding proof of §4, we suppose $k \geqq r + 2$. The case thus omitted is easily handled as was explained just below (4.18).

Proof of (11.6). Using the identity $\Sigma'_{n-2,2} = \Sigma'_{n-3,2} + X_n^{-1}\Sigma'_{n-3,1}$ in (11.6), and then reducing all terms of the resulting relation to a common denominator, we find that (11.6) is equivalent to

$$(11.8) \qquad P\Sigma'_{n-3,\,n-4}(X) > P + \Sigma'_{n-3,\,n-4}(X).$$

Since $n \geqq s + 2 \geqq 5$ and every element in P has a value > 1, each factor in the left member of (11.8) exceeds 2. Consequently (11.8) holds.

Proof of (11.7). It is equivalent to

$$(11.9) \quad \Sigma'_{n-2,\,n-4}(X)\Sigma'_{n-3,\,n-4}(X) - \Sigma'_{n-2,\,n-3}(X)\Sigma'_{n-3,\,n-5}(X) > \Sigma'_{n-3,\,n-5}(X)\,.$$

If $n = 5$, which is the smallest admissible value of n since $n \geqq s + 2$, (11.9) is equivalent to

$$(X_{i_1} + X_{i_2} + X_{i_3})(X_{i_1} + X_{i_2}) - (X_{i_1}X_{i_2} + X_{i_1}X_{i_3} + X_{i_2}X_{i_3}) > 1\,,$$

which is obviously true (with $X_{i_j} > 1$). If $n > 5$, we use the identities

$$\Sigma'_{n-2,\,u}(X) \equiv \Sigma'_{n-3,\,u}(X) + X_n\Sigma'_{n-3,\,u-1}(X) \qquad (u = n - 4, n - 3)\,,$$

and find that (11.9) is equivalent to

$$(11.10) \qquad [\Sigma'_{n-3,\,n-4}(X)]^2 - \Sigma'_{n-3,\,n-3}(X)\Sigma'_{n-3,\,n-5}(X) > \Sigma'_{n-3,\,n-5}(X)\,.$$

Introducing the partition notation $(a^t) = \Sigma'_{n-3,\,t}(X^a)$, we may write (11.10) in the form

$$(11.11) \qquad (1^{n-4})^2 - (1^{n-3})(1^{n-5}) > (1^{n-5})\,.$$

By use of (47) and (48) of I, we find that the left member of (11.11) is not less than

$$(2^{n-4}) + (2^{n-5}1^2)\,,$$

which exceeds (1^{n-5}), since the number of variables under consideration is $n - 3$. Hence (11.7) holds.

Proof of (11.4). Here $n \geqq m + 2 \geqq 6$. Hence we shall obtain the desired result by establishing the inequalities

$$(11.12) \qquad \Sigma'_{n-2,\,m-1} - \Sigma'_{n-3,\,m-1} > \Sigma'_{n-2,\,n-2}(1 + \Sigma'_{n-3,\,1})\,,$$

$$(11.13) \quad (\Sigma'_{n-2,\,m-1}\Sigma'_{n-3,\,t-2} - \Sigma'_{n-2,\,t-2}\Sigma'_{n-3,\,m-1}) > \Sigma'_{n-2,\,n-2}\Sigma'_{n-3,\,t-1} \quad (2 < t \leqq m)\,.$$

Proof of (11.12). Reducing all terms of this relation to a common denominator and using in the result the identity

$$\Sigma'_{n-2,\,n-m-1}(X) \equiv \Sigma'_{n-3,\,n-m-1}(X) + X_n\Sigma'_{n-3,\,n-m-2}(X)\,,$$

we find that (11.12) is equivalent to

$$(11.14) \qquad \Sigma'_{n-3,\,n-m-1}(X) > 1 + P^{-1}\,\Sigma'_{n-3,\,n-4}(X)\,,$$

where P is as in the last paragraph above. No term in the left member of (11.14) is smaller than $X_{i_1} \cdots X_{i_{n-m-1}}$ and no term in $\Sigma'_{n-3,n-4}(X)$ is larger than $X_{i_2} \cdots X_{i_{n-2}}$. Hence (11.14) is true if

$$\binom{n-3}{n-m-1} X_{i_1} \cdots X_{i_{n-m-1}} > 1 + (n-3)/X_{i_1} \qquad (X_{i_j} > 1;\, n \geqq m + 2 \geqq 6)\,. \tag{11.15}$$

The left member of (11.15) is not less than $2(n-3)$ and the right member is not greater than $(n-1)/2$, and so (11.15) holds.

Proof of (11.13). Reducing the terms of (11.13) to a common denominator and employing in the result certain identities analogous to those just above (11.10), we find that (11.13) is equivalent to

$$\Sigma'_{n-3,n-m-1}(X)\Sigma'_{n-3,n-t-1}(X) - \Sigma'_{n-3,n-t}(X)\Sigma'_{n-3,n-m-2}(X) > \Sigma'_{n-3,n-t-2}(X)\,,$$

or, in partition symbols, to

$$(1^{n-m-1})(1^{n-t-1}) - (1^{n-t})(1^{n-m-2}) > (1^{n-t-2}) \qquad (2 < t \leqq m,\, n \geqq m + 2 \geqq 6)\,. \tag{11.16}$$

We let $s_1 = n - t - 1$, $s_2 = n - m - 1$ and use the following relation which can be obtained from (47) and (48) of I:

$$(1^{s_1})(1^{s_2}) - (1^{s_1+1})(1^{s_2-1}) \geqq (2^{s_2}1^{s_1-s_2}) + (2^{s_2-1}1^{s_1-s_2+2}). \tag{11.17}$$

In order to establish (11.16), it suffices to prove that the right member of (11.17) exceeds (1^{s_1-1}). By hypothesis, the number of variables under consideration is $(n-3) > s_1$. Hence neither of the terms in the right member of (11.17) vanishes. Consequently, in the cases where $s_2 = s_1$, $s_2 = s_1 - 1$, and $s_2 = 1$, this right member is readily seen to exceed (1^{s_1-1}). Suppose, then, that $1 < s_2 < s_1 - 1$; we shall prove that

$$(2^{s_2}1^{s_1-s_2}) > (1^{s_1-1}). \tag{11.18}$$

Consider the terms of each of the symbols of (11.18) which involve only the variables of some particular set of s_1 variables (of the $n-3$). The numbers of such terms in the left and right members of (11.18) are $\binom{s_1}{s_2}$ and s_1, respectively. Since $1 < s_2 < s_1 - 1$, $\binom{s_1}{s_2} > s_1$, and every one of the s_1 terms in question of (1^{s_1-1}) is less than the minimum term in question of $(2^{s_2}1^{s_1-s_2})$. By considering in this way every set of s_1 variables from the $n-2$ [and thus counting certain terms of (1^{s_1-1}) more than once, since $n - 3 > s_1$], we find that (11.18), and therefore (11.17), holds.

(ii) $s > r > 1$. We find it desirable to treat this case in the following parts (with $p = n - 1$ in each part): (1) $r = 2$, $s = 3$; (2) $r = 2$, $s = m > 3$; (3) $s > r > 2$.

(1) $r = 2, s = 3$. In this case $\lambda_r = \lambda_2 = 1$, and with $p = n - 1$ inequality (11.1) is

$$\begin{aligned}[t] &[\Sigma'_{n-2,1} - \Sigma'_{n-3,1} + \lambda_3(\Sigma'_{n-2,2} - \Sigma'_{n-3,2}) + \lambda_3^2(\Sigma'_{n-2,2}\Sigma'_{n-3,1} \\ &\qquad - \Sigma'_{n-2,1}\Sigma'_{n-3,2})] > \Sigma'_{n-2,n-2}(\Sigma'_{n-3,1} + \lambda_3\Sigma'_{n-3,2}). \end{aligned} \tag{11.19}$$

To prove (11.19), it suffices to show that (11.7) holds here and that

$$\Sigma'_{n-2,2} - \Sigma'_{n-3,2} > \Sigma'_{n-2,n-2}\Sigma'_{n-3,1}. \tag{11.20}$$

The proof of (11.7) given above applies here. Inequality (11.20) is readily found to be equivalent to $P \equiv X_{i_1} \cdots X_{i_{n-3}} > 1$. This is true since $n - 3 \geqq r = 2$ and $P \geqq X_1 \cdots X_r \geqq 2$.

(2) $r = 2, s = m > 3$. The case $p = n - 1$ of (11.1) now reduces to

$$\begin{aligned} &[\Sigma'_{n-2,1} - \Sigma'_{n-3,1} + \lambda_3(\Sigma'_{n-2,2} - \Sigma'_{n-3,2}) + (\lambda_3^2 - \lambda_4)(\Sigma'_{n-2,2}\Sigma'_{n-3,1} \\ &\quad - \Sigma'_{n-2,1}\Sigma'_{n-3,2}) + \cdots + \lambda_2\lambda_m(\Sigma'_{n-2,m-1} - \Sigma'_{n-3,m-1}) + \lambda_m\lambda_3(\Sigma'_{n-2,m-1}\Sigma'_{n-3,1} \\ &\quad - \Sigma'_{n-2,1}\Sigma'_{n-3,m-1}) + \cdots + \lambda_m^2(\Sigma'_{n-2,m-1}\Sigma'_{n-3,m-2} - \Sigma'_{n-2,m-2}\Sigma'_{n-3,m-1})] \\ &> \Sigma'_{n-2,n-2}(\Sigma'_{n-3,1} + \lambda_3\Sigma'_{n-3,2} + \lambda_4\Sigma'_{n-3,3} + \cdots + \lambda_m\Sigma'_{n-3,m-1}). \end{aligned} \tag{11.21}$$

Since $\lambda_m\lambda_i \geqq \lambda_i$ $(i = 2, \cdots, s)$ and $\lambda_2 = 1$, in order to establish (11.21) it suffices to show that (11.20) and (11.13) hold here, and the proofs of these relations are the same here as they were above.

(3) $s > r > 2$. The case $p = n - 1$ of (11.1) reduces to

$$\begin{aligned} &[A + \lambda_r\lambda_s(\Sigma'_{n-2,s-1}\Sigma'_{n-3,r-2} - \Sigma'_{n-2,r-2}\Sigma'_{n-3,s-1}) + \lambda_{r+1}\lambda_s(\Sigma'_{n-2,s-1}\Sigma'_{n-3,r-1} \\ &\quad - \Sigma'_{n-2,r-1}\Sigma'_{n-3,s-1}) + \cdots + \lambda_s^2(\Sigma'_{n-2,s-1}\Sigma'_{n-3,s-2} - \Sigma'_{n-2,s-2}\Sigma'_{n-3,s-1})] \\ &> \Sigma'_{n-2,n-2}(\Sigma'_{n-3,r-1} + \lambda_{r+1}\Sigma'_{n-3,r} + \lambda_{r+2}\Sigma'_{n-3,r+1} + \cdots + \lambda_s\Sigma'_{n-3,s-1}), \end{aligned} \tag{11.22}$$

where A is the sum of those terms of (11.1) that are obtained here by assigning to i values $> r - 2$ and $< s - 1$. That the sum of the last $s - r + 1$ terms of the left member of (11.22) exceeds the right member of this inequality can be shown by comparing the coefficients of $\lambda_m\lambda_i$ and λ_i in the left and right members, respectively, of (11.22), and observing that (11.13) holds with $m = s$ and $t = r, \cdots, s$ when $r > 2$.

Second step. Let $A_{n-2,r-1}$, $B_{p-2,r-2}$, $A_{n-2,r-2}$, $B_{p-2,r-1}$ be the first, second, third, fourth bracket, respectively, in (11). Since we know that (11) is true for $p = n - 1$, we know that

$$B_{k-2,r-2}(B_{k-2,r-1})^{-1} > A_{n-2,r-2}(A_{n-2,r-1})^{-1} \tag{11.23}$$

is true for $k = n - 1$. Assuming (11.23) to be true for any value of k except the smallest value of p that we admit, namely $p = {}_1q$ if ${}_1q > r$ and $p = {}_1q + 1$

if ${}_1q = r$, we prove that the relation which results when k is replaced by $k - 1$ in (11.23) is also true. In order to do this, it suffices to prove that

$$B_{k-3,r-2}(B_{k-3,r-1})^{-1} \geqq B_{k-2,r-2}(B_{k-2,r-1})^{-1},$$

which is true since the case $p = k - 1$ of (7.1) is a true inequality.

12. **Our most general result relative to equations of the form** (1.5). The methods referred to just before Theorem 8 suffice to prove the following

THEOREM 12. *This is obtained from Theorem 8 by merely substituting in Theorem 8 the $\varphi_n(1/x)$ of (10) for the $\varphi_n(1/x)$ of (6).*

COROLLARY 12. *The Kellogg solution of the equation*

$$\Sigma_{n,1}(1/x) + 2\Sigma_{n,2}(1/x) + 3\Sigma_{n,3}(1/x) + \Sigma_{n,n}(1/x) = 5/11 \quad (n > 4),$$

is $x = w$, where w_i differs from the w_i of Corollary 8 only for $i = n$. Here w_n exceeds the w_n of Corollary 8 by 11.

Part 4. Further results of several different types

13. **Application of the theory to a special equation different in form from both** (1.3) **and** (1.5). In §1, just before describing the new features of the present paper, we indicated the nature of the requirement of consecutive second subscripts of the Σ's in equations (1.3) and (1.5). That one may be able to carry through the procedure for an equation in which n is small and the second subscripts of the Σ's employed are not as we have required them to be generally is shown by the following example.

Consider the equation

$$(13) \qquad \Sigma_{5,2}(1/x) + \Sigma_{5,4}(1/x) = b/a,\ a \equiv [(c+1)b - 1]\,.$$

The Kellogg solution w of (13) can be obtained by properly specializing equation (5).

We indicate our procedure for (13) as follows. Let X be an E-solution of (13) different from w. Suppose that $X_{1\ldots 4}$ contains at least one element of class A and at least one of class B. Suppose also, for the present, that $q_1 = 1$, ${}_1q = 2$. Then let X be transformed into X' in such a way that

$$X_1' = (X_1 - \epsilon) \geqq w_1\,, \qquad X_2' = w_2\,, \qquad X_i' = X_i \qquad (i = 3, 4, 5)\,,$$

while both X and X' satisfy (13). We wish to prove the analog here of the case $f = X$ of (4.8):

$$(13.1) \quad \Sigma_{p,2}(1/X') + \Sigma_{p,4}(1/X') \leqq \Sigma_{p,2}(1/X) + \Sigma_{p,4}(1/X) \qquad (p = 2, 3, 4)\,.$$

According to the definition of the transformation, the case $p = 2$ of (13.1) is

$$(X_1'\, w_2)^{-1} \equiv (X_1'X_2')^{-1} \leqq (X_1X_2)^{-1}\,,$$

whose validity follows from the definition of our transformation and Lemma 1a of I. For $p = 3, 4$ the analogs in question of (4.8) are readily found to be

$$(X_3X_4X_5 + X_3 + X_4 + X_5)(X_3X_4 + X_3X_5 + X_4X_5 + 1)^{-1} \leqq X_3 ,$$

$$(X_3X_4X_5 + X_3 + X_4 + X_5)(X_3X_4 + X_3X_5 + X_4X_5 + 1)^{-1} \leqq (X_3X_4 + 1)(X_3 + X_4)^{-1} ,$$

respectively, and both of these relations are readily found to be true, since X is an E-solution of (13). Had we taken $(q_1, {}_1q)$ to be any other pair of the numbers 1, 2, 3, 4 with $q_1 < {}_1q$, similar results would have been obtained. It now follows that the Kellogg solution w of (13) has the remarkable properties.

14. **A new generalization of Kellogg's problem.**[26] Let a, m be any positive integers and let n be an integer > 1. Then the equation which Kellogg considered is the case $m = a = 1$ of the following equation:

$$\sum_{i=1}^{n} \left(\frac{1}{x_i} + \frac{1}{x_i^2} + \cdots + \frac{1}{x_i^m}\right) = \frac{1}{a} . \tag{14}$$

The Kellogg solution of (14) is $x = w$, where $w_1 = a + 1$,

$$w_{p+1} = aw_1^m \cdots w_p^m + 1 \qquad (p = 1, \cdots , n - 2) ,$$

and w_n is the positive number (greater[27] than w_{n-1}) which satisfies the equation

$$(w_n^{m+1} - w_n^m)\,(w_n^m - 1)^{-1} = a(w_1 \cdots w_{n-1})^m .$$

We state without proof that the methods employed in I for the equation

$$\Sigma_{n,1}(1/x) = b/a , \qquad a = [(c + 1)b - 1] , \tag{14.1}$$

suffice to show here that our solution w of (14) has the remarkable properties.

Remark. If we attempt to find the Kellogg solution, W, of the equation $L = b/a$, where L is the left member of (14) and a is as it is defined in (14.1), we do not find that

$$\sum_{i=1}^{p} \left(\frac{1}{W_i} + \frac{1}{W_i^2} + \cdots + \frac{1}{W_i^m}\right) = \frac{b(W_1 \cdots W_p)^m - 1}{a(W_1 \cdots W_p)^m}$$

for $p = 1$ unless $m = 1$. Consequently we can not apply our usual method of attack in this case. We have found no method that suffices here.

15. **An application of a result in Part 2 to convergence of a type of series.** From the proof that every E-solution of (1.3) as conditioned by (1.4) is a set σ

[26] Cf. O. D. Kellogg, loc. cit., footnote 4.

[27] We leave the proof of this inequality to the reader. It may be established by substitution of w', where $w_i' = w_i$ $(i = 1, \cdots , n - 1)$ and $w_n' = w_{n-1}$ in (14).

[cf. (6.3) and the definition of a set σ in §6], we observe the following fact. Among all infinite series with m-th term equal to

$$U_m = (x_{r+m-1})^{-1} \sum_{j=r}^{r+m-1} \lambda_j \Sigma_{r+m-2,\, j-1}(1/x)\,, \qquad \lambda_r \equiv 1\,, \qquad \lambda_i \equiv 0 \qquad (i > s)\,,$$

where the x's are positive integers such that $U_1 + U_2 + \cdots + U_p < (b/a)$, $a = [(c+1)b - 1]$, for every positive integer p, there is no series which converges to (b/a) more rapidly than does the one that is obtained by letting n increase indefinitely in (1.3) and then taking $x_i = w_i$ $(i = 1, \cdots, n-1)$, where w is the Kellogg solution (5) of (1.3) when the λ's in these equations satisfy (1.4).

16. **The Kellogg solution of the cyclo-symmetric analog of** (1.3). Let n, r, s be integers such that $n \geqq s > r \geqq 1$, and let $C_j\,(1/x)$ stand[28] for the j-th elementary cyclo-symmetric function of the n reciprocals $1/x_1, \cdots, 1/x_n$. We consider the equation

$$\begin{aligned} C_r(1/x) + \lambda_{r+1}C_{r+1}(1/x) + \lambda_{r+2}C_{r+2}(1/x) + \cdots + \lambda_s C_s(1/x) &= b/a\,, \\ a &= [(c+1)\,b - 1]\,, \end{aligned} \tag{16}$$

in which the λ's are integers $\geqq 0$. The Kellogg solution of (16) is $x = w$, where

$$\left\{\begin{aligned} w_p &= 1 && (p = 1, \cdots, r-1)\,; \\ w_r &= c + 1\,; \\ w_p &= \sum_{j=r}^{p} a\lambda_j w_1 \cdots w_{p-j} + 1 && (p = r+1, \cdots, n-1)\,; \\ w_n &= \sum_{j=r}^{s} \sum_{i=1}^{j} a\lambda_j w_i w_{i+1} \cdots w_{n-1-j+i}\,, \end{aligned}\right. \tag{16.1}$$

and $\lambda_r \equiv 1$, $\lambda_j \equiv 0$ $(s < j < n)$, $w_k \cdots w_{k-1} \equiv 1$.

The solution w defined by (16.1) has certain properties which for the case $\lambda_i = 0$ $(i = r+1, \cdots, s)$ were discussed in the article referred to in footnote 28. We have not been able to prove or disprove that this solution has the remarkable properties relative to equation (16).

NORTHWESTERN UNIVERSITY.

[28] The function $C_r(1/x)$ was the left member of the equation considered by Simmons in an article *On a cyclo-symmetric equation*, American Mathematical Monthly, vol. 36 (1929), pp. 148–155.

λ-DEFINABILITY AND RECURSIVENESS

By S. C. Kleene

1. Introduction. In Kleene [2][1] a theory of the definition of functions of positive integers by certain formal means is developed in connection with the study of a system of formal logic.[2] The system of formal logic is shown in Kleene-Rosser [1] to be inconsistent; however, the theory of formal definition remains of interest, both for its use in a new system of formal logic proposed by Church in [3], and for its connection with questions of constructibility and decidability in number theory.[3] Hence it seems desirable to bring together the essentials of the theory, and to develop them from a somewhat new point of view, in which the emphasis is on the connection with the recursive functions. In this presentation, no knowledge of systems of formal logic is presupposed, but use will be made of a few results of the intuitive theory of recursive functions.[4]

It is found convenient here to treat the functions as functions of natural numbers, rather than of positive integers. This change can be regarded as a change merely in the notation.

The theory deals with a class of formulas composed of the symbols {, }, (,), λ, [,] and other symbols f, x, ρ, $\cdots$ called variables or *proper symbols*, where f, x, ρ, $\cdots$ is a given infinite list.

A formula is called *properly-formed* if it is obtainable from proper symbols by zero or more successive operations of combining **M** and **N** to form {**M**} (**N**) or λ**x**[**M**], where **x** is any proper symbol. An occurrence of a proper symbol **x** in a formula **F** is called *bound* or *free* according as it is or is not an occurrence in a properly-formed part of the form λ**x**[**M**]. By a free (bound) symbol of **F** is meant a proper symbol which occurs in **F** as a free (bound) symbol. A formula shall be *well-formed*, if it is properly-formed, and if, for each properly-formed part of the form λ**x**[**M**], where **x** is a proper symbol, **x** is a free symbol of **M**.

Heavy-typed letters will henceforth represent undetermined well-formed formulas under the convention that each set of symbols standing apart in which a heavy-typed letter occurs shall stand for a well-formed formula.[5] As abbre-

Received July 1, 1935; presented to the American Mathematical Society, September 13, 1935.

[1] The numbers in brackets refer to the bibliography at the end.

[2] Use is made, directly or indirectly, of Church [1]-[2], Kleene [1], Rosser [1], Curry [1]-[3], Schönfinkel [1].

[3] See Kleene [2] p. 232, Church [4], and Church-Kleene [1].

[4] In writing this paper, I have profited from discussion of the subject with Dr. J. B. Rosser, and I also thank him for assistance with the manuscript.

[5] A detailed analysis of the structure of well-formed formulas, and of the implications of this convention, is given in Kleene [1] §§2, 3. The term "proper symbol" was introduced in place of "variable" in order to save the latter for use in another meaning in connection with the formal logics under consideration.

viations, we shall write $\{\mathbf{F}\}(\mathbf{A}_1, \cdots, \mathbf{A}_n)$ or $\mathbf{F}(\mathbf{A}_1, \cdots, \mathbf{A}_n)$ instead of $\{\cdots\{\mathbf{F}\}(\mathbf{A}_1)\cdots\}(\mathbf{A}_n)$, and $\lambda\mathbf{x}_1 \cdots \mathbf{x}_n\cdot\mathbf{M}$ instead of $\lambda\mathbf{x}_1[\cdots\lambda\mathbf{x}_n[\mathbf{M}]\cdots]$. $S^{\mathbf{x}_1 \cdots \mathbf{x}_n}_{\mathbf{A}_1 \cdots \mathbf{A}_n}\mathbf{M}\,|$ shall denote the result of substituting $\mathbf{A}_i$ for each of the occurrences (if any) of $\mathbf{x}_i$ in $\mathbf{M}$ $(i = 1, \cdots, n)$. From time to time we assign individual symbols to stand as abbreviations for particular formulas, indicating this by an arrow $\rightarrow$, as

$$I \rightarrow \lambda f\cdot f\,, \qquad\qquad J \rightarrow \lambda fxyz\cdot f(x, f(z, y))\,.$$

We introduce an equivalence relation **A** conv **B**, or **A** is *convertible* into **B**, between well-formed formulas, which is defined to be the relation of least domain which is (1) reflexive, (2) symmetric, and (3) transitive, and has further the properties (4) if **A** conv **B**, then $\{\mathbf{C}\}(\mathbf{A})$ conv $\{\mathbf{C}\}(\mathbf{B})$, $\{\mathbf{A}\}(\mathbf{C})$ conv $\{\mathbf{B}\}(\mathbf{C})$, and $\lambda\mathbf{x}[\mathbf{A}]$ conv $\lambda\mathbf{x}[\mathbf{B}]$, (5) if the proper symbol **y** does not occur in **A**, $\lambda\mathbf{x}[\mathbf{A}]$ conv $S^{\mathbf{x}}_{\mathbf{y}}\lambda\mathbf{x}[\mathbf{A}]\,|$, and (6) if **x** and the free symbols of **N** are not bound symbols of **M**, $\{\lambda\mathbf{x}[\mathbf{M}]\}(\mathbf{N})$ conv $S^{\mathbf{x}}_{\mathbf{N}}\mathbf{M}\,|$.[6]

If $\{\mathbf{F}\}(\mathbf{N})$ is interpreted as representing the value of **F** (considered as a function) for **N** as argument, and $\lambda\mathbf{x}[\mathbf{M}]$ as representing the function which **M** is of **x**, then the equivalence relation **A** conv **B** corresponds to a relation of equality in meaning.[7] The analysis of the relation **A** conv **B** given in Church-Rosser [1] can be regarded as furnishing a demonstration of the consistency of the system under these interpretations: A formula $\overline{\mathbf{A}}$ which has no part of the form $\{\lambda\mathbf{x}[\mathbf{M}]\}(\mathbf{N})$ is said to be in *normal form*, and to be a normal form of any formula **A** convertible into it. According to Theorem 1, Corollary 2, if **A** has a normal form, the latter is unique to within the choice of symbols used in it as bound variables.[8]

Evidently, a demonstration that **A** conv **B** is given by passing from **A** to **B** by successive substitutions (on individual parts of a formula not immediately following λ) of (a) **C** for **D** (or inversely), where **D** conv **C** is known, and (b) $S^{\mathbf{x}}_{\mathbf{N}}\mathbf{M}\,|$ for $\{\lambda\mathbf{x}\cdot\mathbf{M}\}(\mathbf{N})$ (or inversely), changing bound variables when necessary to avoid confounding variables that should be distinct or to reach a desired formula.

The substitution $S^{\mathbf{x}_1 \cdots \mathbf{x}_n}_{\mathbf{N}_1 \cdots \mathbf{N}_n}\mathbf{M}\,|$ for $\{\lambda\mathbf{x}_1\cdots\mathbf{x}_n\cdot\mathbf{M}\}(\mathbf{N}_1, \cdots, \mathbf{N}_n)$ is equivalent to a series of the substitutions (b). Indeed, from the interpretations of $\{\mathbf{F}\}(\mathbf{N})$ and $\lambda\mathbf{x}[\mathbf{M}]$, it follows that the expression which we abbreviate to $\mathbf{F}(\mathbf{N}_1, \cdots, \mathbf{N}_n)$ represents the value of **F** (considered as a function of n variables) for the set of

[6] (1) and the clause "$\{\mathbf{A}\}(\mathbf{C})$ conv $\{\mathbf{B}\}(\mathbf{C})$" of (4) are redundant. The present definition is equivalent to the former one, according to which **A** conv **B** whenever **B** is derivable from **A** by certain rules I–III, the derivation being called a *conversion* (cf. both Church [1] and Kleene [1]).

[7] *A* relation, rather than *the* relation, since, for example, it can be maintained that $\lambda fx\cdot f(x)$ and $\lambda f\cdot f$ have the same meaning.

[8] The notion of the *normal form* of a formula under conversion was originally introduced by Church in lectures at Princeton in the fall of 1931.

arguments $\mathbf{N}_1, \cdots, \mathbf{N}_n$; and the expression which we abbreviate $\lambda \mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{M}$ represents the function which $\mathbf{M}$ is of $\mathbf{x}_1, \cdots, \mathbf{x}_n$.[9]

We have specified a class of formulas (the well-formed formulas) and an equivalence relation between formulas of this class (the relation of interconvertibility). We bring the natural numbers into relation with this subject-matter by selecting a progression of well-formed formulas

$$\lambda fx \cdot f(x), \qquad \lambda fx \cdot f(f(x)), \qquad \lambda fx \cdot f(f(f(x))), \cdots$$

to "represent" or "be identified with" the natural numbers in our symbolism. This is recognized in the notation by assigning

$$0 \rightarrow \lambda fx \cdot f(x), \qquad 1 \rightarrow \lambda fx \cdot f(f(x)), \qquad 2 \rightarrow \lambda fx \cdot f(f(f(x))), \cdots.$$

It may now happen, for a non-negative integral function $L(x_1, \cdots, x_n)$ of natural numbers, that there are well-formed formulas $\mathbf{L}$ which automatically represent the function $L(x_1, \cdots, x_n)$, on the basis of our equivalence relation and our interpretation of $\mathbf{F}(\mathbf{N}_1, \cdots, \mathbf{N}_n)$. That is, there may be formulas $\mathbf{L}$ such that, whenever $\mathbf{x}_1, \cdots, \mathbf{x}_n$ represent natural numbers $x_1, \cdots, x_n$, respectively, $\mathbf{L}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ is convertible into the formula which represents the natural number $L(x_1, \cdots, x_n)$. In this case, we shall say that $L(x_1, \cdots, x_n)$ is (*formally*) *defined* or λ-*defined* by $\mathbf{L}$.

Thus a problem arises: what functions $L(x_1, \cdots, x_n)$ are λ-definable? We have at once that the successor function is λ-definable, since

$$\{\lambda \rho fx \cdot f(\rho(f, x))\} \ (\lambda fx \cdot f(\cdots n+1 \text{ times} \cdots f(x) \cdots)) \text{ conv } \lambda fx \cdot f(\cdots n+2 \text{ times} \cdots f(x) \cdots) \quad (n = 0, 1, 2, \cdots).$$

Accordingly let

$$S \rightarrow \lambda \rho fx \cdot f(\rho(f, x)).$$

The identity function of a natural number is also λ-definable, since the formula which we have called I has the property $I(\mathbf{x})$ conv $\mathbf{x}$.

The problem has arisen from the point of view in which interconvertible formulas are regarded as equivalent. Hence we should consider whether the representations involved are unambiguous from this point of view. Let us call a representation of a class of mathematical entities by well-formed formulas *well-founded* if interconvertible formulas cannot represent different entities of the class. It follows from the above-mentioned consistency proof (Church-Rosser [1]) that the given representation of natural numbers by well-formed formulas is well-founded; this in turn implies that such representation of non-negative integral functions of natural numbers as λ-definition yields is well-founded.[10]

[9] This device for expressing functions of several variables in terms of functions for one variable goes back to Schönfinkel [1].

[10] Non-interconvertible formulas may represent the same entity of a given class, and a formula may represent entities of different classes, e.g., the formulas abbreviated I and 0 both represent the identity function of a natural number, while the latter also represents the natural number 0.

The problem is a special case of the larger problem: what functional relationships among well-formed expressions can be expressed by well-formed formulas?

We shall say, generally, that a function L which associates well-formed formulas with finite ordered sets of well-formed formulas is (*formally*) *defined* or λ-*defined* by **L** if for each finite ordered set $\mathbf{A}_1, \cdots, \mathbf{A}_{n_{\mathbf{A}}}$ for which L is defined, $\mathbf{L}(\mathbf{A}_1, \cdots, \mathbf{A}_{n_{\mathbf{A}}})$ is convertible into the value of the function L for the set $\mathbf{A}_1, \cdots, \mathbf{A}_{n_{\mathbf{A}}}$ of arguments; and we shall understand, by the λ-definition of a function of which the arguments or the values are other mathematical entities, the λ-definition of a function which corresponds under the representation of the mathematical entities by well-formed formulas (in case a representation has been specified).[11]

In this paper we restrict ourselves (except incidentally) to the case of the larger problem in which the independent variables are fixed in number, and range over the natural numbers. The subcase of the problem in which the values are also natural numbers (i.e., the problem first proposed) we treat in §§2, 3 by proving that all recursive functions, in a wide sense of the term recursive, due to Herbrand and Gödel, are λ-definable; and conversely, all λ-definable functions of the type in question are recursive. In §§4, 5 it is shown that, using the term recursive in an extended sense, these results can be generalized (under additional hypotheses) to the case in which the values are any well-formed expressions. The extended sense of the term recursive is obtained by assigning numbers to the values, by the Gödel method, and requiring that they be a recursive function of the arguments in the first sense.

The formulas $\lambda fx \cdot f(x)$, $\lambda fx \cdot f(f(x))$, $\cdots$ were originally coördinated with the positive integers 1, 2, $\cdots$ (Church [2], p. 863). That is a suitable course to follow in developing number-theory (Kleene [2]). In this paper, for technical reasons, we are using instead the correspondence established above between those formulas and the natural numbers 0, 1, $\cdots$. Because of this, the concept of λ-definability of a function is altered. But, for the interpretation of the final results, one can easily go back to the original notion of λ-definability. Since the "natural numbers", "0", "1", $\cdots$ enter into our definitions of λ-definable function and recursive function in the rôle of a progression, it is only necessary to rename them "positive integers", "1", "2", $\cdots$ in those definitions. Or one may use the following relation: A positive integral function $\phi(y_1, \cdots, y_n)$ [well-formed function $\Phi(y_1, \cdots, y_n)$] of positive integers $y_1, \cdots, y_n$ is λ-definable in the original sense if and only if the function $\phi(x_1 + 1, \cdots, x_n + 1) - 1$ $[\Phi(x_1 + 1, \cdots, x_n + 1)]$ of natural numbers $x_1, \cdots, x_n$ is λ-definable in the present sense.[12]

[11] The λ-definition of a sequence $\mathbf{A}_0, \mathbf{A}_1, \cdots$ shall mean the λ-definition of the function which $\mathbf{A}_i$ is of i, and the λ-enumeration of a class shall mean the λ-definition of an enumeration (with or without repetitions) of the members of the class.

[12] Under the Church representation of the positive integers, the formula of n denotes the operation of applying the n-th power of a function to an argument, and exceedingly simple

2. **Recursive non-negative integral functions.** In the λ-notation, the definition of a function by substitution is immediate:

(1) $\{\lambda \mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{G}(\mathbf{H}_1(\mathbf{x}_1, \cdots, \mathbf{x}_n), \cdots, \mathbf{H}_m(\mathbf{x}_1, \cdots, \mathbf{x}_n))\}(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ *conv* $\mathbf{G}(\mathbf{H}_1(\mathbf{X}_1, \cdots, \mathbf{X}_n), \cdots, \mathbf{H}_m(\mathbf{X}_1, \cdots, \mathbf{X}_n))$.[13]

When an italic letter denotes a number, the same letter in heavy type shall denote the corresponding formula. Our remark that S λ-defines the successor function of a natural number can now be written thus:

(2) *If* $x + 1 = z$, $S(\mathbf{x})$ *conv* $\mathbf{z}$ $(x = 0, 1, \cdots)$.

In view of the form of $\mathbf{x}$ $(x = 0, 1, \cdots)$:

(3) $\mathbf{x}(\mathbf{F}, \mathbf{A})$ *conv* $\mathbf{F}(\cdots x + 1 \text{ times} \cdots \mathbf{F}(\mathbf{A}) \cdots)$ $\qquad (x = 0, 1, \cdots)$.

(4) $\mathbf{x}(I)$ *conv* I $\quad (x = 0, 1, \cdots)$. $\quad (I(\mathbf{A})$ *conv* $\mathbf{A}.)$

By use of (4):

(5) $\{\lambda t \cdot t(I, 0)\}(\mathbf{x})$ *conv* 0 $\qquad (x = 0, 1, \cdots)$.

(6) $\{\lambda t_1 \cdots t_n \cdot t_1(I, \cdots, t_n(I, t_i) \cdots)\}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ *conv* $\mathbf{x}_i$ $\quad (x_1, \cdots, x_n = 0, 1, \cdots)$.

λ-definitions of addition, multiplication, and exponentiation, due to Rosser (Kleene [2] pp. 160–164), are possible.

If that representation is extended by adding $\lambda f x \cdot x(f)$ to represent 0, the resulting formal definition of functions of natural numbers is equivalent to the one of this paper in respect to the results we have summarized.

If the Church representation is extended by the natural method of using the class of properly-formed formulas instead of the class of well-formed formulas, modifying suitably the relation *conv*, and letting $\lambda f x \cdot x$ represent 0, simplifications are afforded in the proofs of many theorems, but unfortunately difficulties are introduced in the formal logics in which this theory is used. Rosser has shown that the formal definition (λ-K-definition) under this program is equivalent to λ-definition, when the range of the independent variables is the set of natural numbers, and all the values have the same free symbols. For functions over all well-formed formulas, λ-K-definition is not equivalent to λ-definition, but we conjecture that the equivalence holds for many other significant ranges of the independent variables (such as functions of natural numbers, functions of functions of natural numbers, $\cdots$, with values in the same range, and ordinal numbers represented by well-formed formulas as in Church-Kleene [1]), and fails only for very heterogeneous ranges.

The formal definition which is obtained from that of this paper by using the [λ-δ-]conversion of Church [3] is likewise equivalent to λ-definition, when the range of the independent variables is the set of natural numbers and the values do not contain δ.

[13] Here we assume that $\mathbf{x}_1, \cdots, \mathbf{x}_n$ do not occur in $\mathbf{G}, \mathbf{H}_1, \cdots, \mathbf{H}_m$ as free symbols; and, in general, when a heavy-typed letter represents occurrences of a proper symbol in a formula, we suppose the only occurrences of the symbol in the formula to be those appearing explicitly, unless the contrary is implied by the original convention concerning heavy-type.

For the moment we abbreviate $\{\lambda\rho\sigma\tau fgha\cdot\rho(f, \sigma(g, \tau(h, a)))\}(\mathbf{x}, \mathbf{y}, \mathbf{z})$ to $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$, $\{\lambda\rho f\cdot\rho(f, I, I)\}(\mathbf{X})$ to $\mathbf{X}_1$, $\{\lambda\rho f\cdot\rho(I, f, I)\}(\mathbf{X})$ to $\mathbf{X}_2$, $\{\lambda\rho f\cdot\rho(I, I, f)\}(\mathbf{X})$ to $\mathbf{X}_3$.

(7) $[\mathbf{x}, \mathbf{y}, \mathbf{z}]_1$ *conv* $\mathbf{x}$, $[\mathbf{x}, \mathbf{y}, \mathbf{z}]_2$ *conv* $\mathbf{y}$, $[\mathbf{x}, \mathbf{y}, \mathbf{z}]_3$ *conv* $\mathbf{z}$ $(x, y, z = 0, 1, \cdots)$.

If $\mathfrak{F} \rightarrow \lambda\rho\cdot[\rho_2, \rho_3, S(\rho_3)]$ and $\mathfrak{I} \rightarrow [0, 0, 0]$, then, using (2) and (7):

(8) $\mathfrak{F}(\mathfrak{I})$ *conv* $[0, 0, 1]$, $\mathfrak{F}(\mathfrak{F}(\mathfrak{I}))$ *conv* $[0, 1, 2]$, $\mathfrak{F}(\mathfrak{F}(\mathfrak{F}(\mathfrak{I})))$ *conv* $[1, 2, 3]$, $\mathfrak{F}(\mathfrak{F}(\mathfrak{F}(\mathfrak{F}(\mathfrak{I}))))$ *conv* $[2, 3, 4]$, $\cdots$.

Hence, letting $P \rightarrow \lambda\rho\cdot\rho(\mathfrak{F}, \mathfrak{I})_1$, and using (3), (7) and (8):

(9) *If* $x = z + 1$, $P(\mathbf{x})$ *conv* $\mathbf{z}$ $(x = 1, 2, \cdots)$. $P(0)$ *conv* 0.

Now let $\dot{-} \rightarrow \lambda\mu\nu\cdot\nu(P, S(\mu))$, and abbreviate $\{\dot{-}\}(\mathbf{x}, \mathbf{y})$ to $[\mathbf{x}] \dot{-} [\mathbf{y}]$ (omitting brackets when no ambiguity results). By (3) and (9):

(10) *If* $x \geqq y$ *and* $x - y = z$, $\mathbf{x} \dot{-} \mathbf{y}$ *conv* $\mathbf{z}$; *if* $x \leqq y$, $\mathbf{x} \dot{-} \mathbf{y}$ *conv* 0 $(x, y = 0, 1, \cdots)$.

Let min $\rightarrow \lambda xy\cdot y \dot{-} [y \dot{-} x]$.

(11) *If* $x \leqq y$, min $(\mathbf{x}, \mathbf{y})$ *conv* min $(\mathbf{y}, \mathbf{x})$ *conv* $\mathbf{x}$ $(x, y = 0, 1, \cdots)$.

We call a formula constructed out of I, J and proper symbols by zero or more successive operations of passing from $\mathbf{F}$ and $\mathbf{A}$ to $\{\mathbf{F}\}(\mathbf{A})$ a *combination*, and the individual occurrences in it of I, J and proper symbols which enter in the course of this construction its *terms*. Let $T \rightarrow J(I, I)$, so that T conv $\lambda fx\cdot x(f)$. The reader may verify that

(12) $J(T, \mathbf{A}, \mathbf{F})$ *conv* $\lambda\mathbf{x}\cdot\mathbf{F}(\mathbf{x}, \mathbf{A})$, $J(T, \mathbf{A}, J(I, \mathbf{F}))$ *conv* $\lambda\mathbf{x}\cdot\mathbf{F}(\mathbf{A}(\mathbf{x}))$, $J(T, T, J(I, J(T, T, J(T, A, J(T, \mathbf{F}, J)))))$ *conv* $\lambda\mathbf{x}\cdot\mathbf{F}(\mathbf{x}, \mathbf{A}(\mathbf{x}))$.

If $\mathbf{C}$ is the proper symbol $\mathbf{x}$, I is a combination convertible into $\lambda\mathbf{x}\cdot\mathbf{C}$. If $\mathbf{C}$ is a combination of the form $\mathbf{F}(\mathbf{A})$ and has $\mathbf{x}$ as a free symbol, then $\mathbf{x}$ is a free symbol either (a) of $\mathbf{F}$ but not of $\mathbf{A}$, (b) of $\mathbf{A}$ but not of $\mathbf{F}$, or (c) both of $\mathbf{F}$ and of $\mathbf{A}$. In case (a), if $\mathbf{F}^\circ$ is a combination convertible into $\lambda\mathbf{x}\cdot\mathbf{F}$, then by (12) $J(T, \mathbf{A}, \mathbf{F}^\circ)$ is a combination convertible into $\lambda\mathbf{x}\cdot\mathbf{F}^\circ(\mathbf{x}, \mathbf{A})$ and hence into $\lambda\mathbf{x}\cdot\mathbf{F}(\mathbf{A})$ or $\lambda\mathbf{x}\cdot\mathbf{C}$, and similarly in cases (b) and (c). Thus, by induction on the number of terms of $\mathbf{C}$:

(13) *If* x *is a free symbol of the combination* $\mathbf{C}$, *there is a combination* $\mathbf{C}^\circ$ *such that* $\mathbf{C}^\circ$ *conv* $\lambda\mathbf{x}\cdot\mathbf{C}$.

A proper symbol is a combination; if $\mathbf{F}'$ and $\mathbf{A}'$ are combinations convertible into $\mathbf{F}$ and $\mathbf{A}$, respectively, $\mathbf{F}'$ $(\mathbf{A}')$ is a combination convertible into $\mathbf{F}(\mathbf{A})$; if $\mathbf{R}$ has $\mathbf{x}$ as a free symbol, and $\mathbf{R}'$ is a combination convertible into $\mathbf{R}$, then $\mathbf{R}'$ has $\mathbf{x}$ as a free symbol (interconvertible formulas have the same free symbols), and by (13) there is a combination $\mathbf{R}'^\circ$ convertible into $\lambda\mathbf{x}\cdot\mathbf{R}'$ and hence into $\lambda\mathbf{x}\cdot\mathbf{R}$.

Thus, by an induction corresponding to the process of construction of a well-formed formula:

(14) *Given* $\mathbf{A}$, *there is a combination* $\mathbf{A}'$ *such that* $\mathbf{A}'$ *conv* $\mathbf{A}$.[14]

Let $H \rightarrow \lambda\sigma \cdot \sigma(I, I, I, I)$. Given a formula $\mathbf{A}$ having no free symbols, there is by (14) a combination $\mathbf{A}'$ convertible into $\mathbf{A}$. $\mathbf{A}'$ has no free symbols, and hence its terms are I's and J's. Let $\mathrm{S}'^{\mathrm{T}}_{y(\mathrm{T})}\,\mathbf{A}'|$ denote the result of replacing each term T of $\mathbf{A}'$ by $y(\mathrm{T})$, and let $\mathbf{C} \rightarrow \lambda y \cdot \mathrm{S}'^{\mathrm{T}}_{y(\mathrm{T})}\,\mathbf{A}'|$. Then $\mathbf{C}(I)$ conv $\mathrm{S}'^{\mathrm{T}}_{I(\mathrm{T})}\,\mathbf{A}'|$ conv $\mathbf{A}'$ conv $\mathbf{A}$; and $\mathbf{C}(H)$ conv $\mathrm{S}'^{\mathrm{T}}_{H(\mathrm{T})}\,\mathbf{A}'|$ conv $\mathrm{S}'^{\mathrm{T}}_{I}\,\mathbf{A}'|$ conv I.

(15) *If* $\mathbf{A}$ *has no free symbols, there is a formula* $\mathbf{C}$ *such that* $\mathbf{C}(I)$ *conv* $\mathbf{A}$ *and* $\mathbf{C}(H)$ *conv* I.

Let $B_{-1} \rightarrow \lambda pxyz \cdot p(x, z, y)$, $B_0 \rightarrow \lambda pxyz \cdot p(y, x, z)$, $B_1 \rightarrow \lambda pxyz \cdot p(x, y, z)$.

(16) $B_1(B_{-1})$ *conv* B_{-1}, $B_{-1}(B_1)$ *conv* B_0, $B_{-1}(B_{-1}(B_1))$ *conv* B_1.

We now adopt the notation $-1 \rightarrow \lambda fx \cdot x(f)$. Given formulas $\mathbf{A}_{-1}, \mathbf{A}_0, \mathbf{A}_1$ having no free symbols, there are by (15) formulas $\mathbf{C}_i$ such that $\mathbf{C}_i(I)$ conv $\mathbf{A}_i$ and $\mathbf{C}_i(H)$ conv I $(i = -1, 0, 1)$. Then $\lambda n \cdot n(B_{-1}, B_1, \lambda abc \cdot b(a(H, c(H))), \mathbf{C}_0, \mathbf{C}_1, \mathbf{C}_{-1})$ has the properties of $\mathbf{F}$ in the following:

(17) *If* $\mathbf{A}_{-1}, \mathbf{A}_0, \mathbf{A}_1$ *have no free symbols, there is a formula* $\mathbf{F}$ *such that* $\mathbf{F}(\mathbf{i})$ *conv* $\mathbf{A}_i$ $(i = -1, 0, 1)$.

If $\mathfrak{A} \rightarrow \lambda\rho \cdot \rho(0, 1)$, then, using (3) and the relations $0(1)$ conv 1 and $1(0)$ conv 0:

(18) $\mathfrak{A}(\mathbf{n})$ *conv* 1 $(n = 0, 1, \cdots)$. $\mathfrak{A}(-1)$ *conv* 0.

If $\mathbf{F}$ has no free symbols, there is by (17) a formula $\mathbf{B}$ such that $\mathbf{B}(-1)$ conv $\mathbf{B}(0)$ conv I and $\mathbf{B}(1)$ conv $\lambda bx \cdot \mathbf{F}(x, \lambda\rho \cdot b(\mathfrak{A}(\rho), b, \rho))$. Then $\lambda\rho \cdot \mathbf{B}(\mathfrak{A}(\rho), \mathbf{B}, \rho)$ has the properties of $\mathbf{L}$ in the following:

(19) *If* $\mathbf{F}$ *has no free symbols, there is a formula* $\mathbf{L}$ *such that* $\mathbf{L}(\mathbf{x})$ *conv* $\mathbf{F}(\mathbf{x}, \mathbf{L})$ $(x = 0, 1, \cdots)$ *and* $\mathbf{L}(-1)$ *conv* I.

Given formulas $\mathbf{G}$ and $\mathbf{H}$ having no free symbols, choose $\mathbf{K}$ by (17) so that $\mathbf{K}(0)$ conv $\lambda yf \cdot y(f(-1), \mathbf{G})$ and $\mathbf{K}(1)$ conv $\lambda yfx_2 \cdots x_n \cdot \mathbf{H}(P(y), f(P(y), x_2, \cdots, x_n), x_2, \cdots, x_n)$, and let $\mathbf{F} \rightarrow \lambda y \cdot \mathbf{K}(\min(y, 1), y)$. Then the $\mathbf{L}$ given by (19) for this $\mathbf{F}$ satisfies the following:

(20) *If* $\mathbf{G}$ *and* $\mathbf{H}$ *have no free symbols, there is a formula* $\mathbf{L}$ *such that*

$$\mathbf{L}(0, \mathbf{x}_2, \cdots, \mathbf{x}_n) \textit{ conv } \mathbf{G}(\mathbf{x}_2, \cdots, \mathbf{x}_n) \textit{ and } \mathbf{L}(S(\mathbf{y}), \mathbf{x}_2, \cdots, \mathbf{x}_n) \textit{ conv}$$

$$\mathbf{H}(\mathbf{y}, \mathbf{L}(\mathbf{y}, \mathbf{x}_2, \cdots, \mathbf{x}_n), \mathbf{x}_2, \cdots, \mathbf{x}_n) \; (y, x_2, \cdots, x_n = 0, 1, \cdots).$$

Choose $\mathbf{K}$ by (17) so that $\mathbf{K}(0)$ conv $\lambda fyr \cdot r(y, f(-1), y)$ and $\mathbf{K}(1)$ conv $\lambda fyr \cdot f(r(S(y)), S(y), r)$, let $\mathbf{F} \rightarrow \lambda x \cdot \mathbf{K}(\min(x, 1))$, and choose $\mathbf{L}$ by (19) for this $\mathbf{F}$. Then $\mathbf{L}(\mathbf{x}, \mathbf{y}, \mathbf{r})$ conv $\mathbf{L}(\mathbf{r}(S(\mathbf{y})), S(\mathbf{y}), \mathbf{r})$ $(x = 1, 2, \cdots)$ and, if $\mathbf{r}(\mathbf{y})$ conv

[14] This theorem derives from Rosser [1], and the present proof of it from Church [3].

z, where z is a natural number, **L**(0, **y**, **r**) conv **y**. Hence, letting $\mathfrak{e}_n \to \lambda r x_1 \cdots x_n \cdot \mathbf{L}(r(x_1, \cdots, x_n, 0), 0, r(x_1, \cdots, x_n))$:

(21) *If* **r** *λ-defines a non-negative integral function* $\rho(x_1, \cdots, x_n, y)$ *of natural numbers such that* $(x_1, \cdots, x_n)(Ey)[\rho(x_1, \cdots, x_n, y) = 0]$, then $\mathfrak{e}_n(\mathbf{r})$ *λ-defines* $\epsilon y[\rho(x_1, \cdots, x_n, y) = 0]$.[15]

According to Kleene [3] IV every function of natural numbers recursive in the general Herbrand-Gödel sense (see [3] Def. 2a or Def. 2b) is expressible in the form $\psi(\epsilon y[\rho(x_1, \cdots, x_n, y) = 0])$, where $\psi(y)$ and $\rho(x_1, \cdots, x_n, y)$ are primitive recursive ([3] Def. 1) and $(x_1, \cdots, x_n)(Ey)[\rho(x_1, \cdots, x_n, y) = 0]$. In view of (1), (2), (5), (6) and (20),[16] every primitive recursive function is λ-definable; and therefore, from (21), every general recursive function is λ-definable.

(22) *Every non-negative integral function of natural numbers which is recursive in the Herbrand-Gödel sense is λ-definable.*

(19) constitutes a schema for circular definition. Given any set of conditions of dependence of an entity $L(x)$ on the variable natural number x and on L itself, if the set can be expressed in the λ-notation by a formula **F**, a formula **L** satisfying the conditions in terms of the equivalence relation **A** conv **B** can be found.[17] To do this it need not be known that the conditions actually determine a function $L(x)$. Further analysis of this situation (Kleene [2] §18) shows that to each problem of a large class, which includes many famous unsolved problems (such as the Fermat problem and the 4-color problem), there is a formula **P** such that *whether* **P** *has a normal form* is an equivalent problem.

3. **λ-definable non-negative integral functions.** We now set up a representation of the well-formed formulas by natural numbers, by the Gödel method. The symbols which occur in well-formed formulas we number thus:[18]

$$\lambda \ldots 1;\ \{, (, [\ \ldots 11;\ \},),]\ \ldots 13; \text{ the } i\text{-th proper symbol} \ldots p_{i+6}$$

(p_i = the i-th prime number), and we order numbers to formulas (considered as finite sequences of symbols), finite sequences of formulas, etc., on the basis of the correspondence $n_1, n_2, \cdots, n_k$ to $p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$ between finite sequences of numbers and individual numbers ($n_1, n_2, \cdots, n_k > 0$). Using the methods

[15] Read $(x_1, \cdots, x_n)$ "for all $x_1, \cdots, x_n$", (Ey) "there is a y", $\epsilon y[R(y)]$ "the least y such that $R(y)$ (0 if there is no such y)".

[16] A formula which λ-defines a non-negative integral function of natural numbers has no free symbols.

[17] (17) can be used in the selection of **F**, if cases are distinguished in the form of the dependence of $L(x)$ on x and L. (4) and the clause $\mathbf{L}(-1)$ conv I of (19) can be used whenever under a given case $L(x)$ is independent of either or both of x and L. These devices are illustrated in the proofs of (20), (21) and (24).

[18] The distinction in well-formed formulas of three species of parentheses is unessential, since the species of each parenthesis can be determined from its situation.

and notations of Gödel [1] pp. 179–182,[19] and starting from 1–5, 7–10 of his list (p. 182) and 6, 11–18 of Kleene [3], we define the additional primitive recursive functions and relations 19–42:

19. $Z(0) = 2 \cdot 3^{17} \cdot 5^{11} \cdot 7 \cdot 11^{19} \cdot 13^{11} \cdot 17^{11} \cdot 19^{17} \cdot 23^{13} \cdot 29^{11} \cdot 31^{19} \cdot 37^{13} \cdot 41^{13} \cdot 43^{13}$,

$$Z(n+1) = Su\, Z(n) \begin{pmatrix} 11 + 4n \\ 2^{11} \cdot 3^{17} \cdot 5^{13} \cdot 7^{11} \cdot 11^{19} \cdot 13^{13} \end{pmatrix}.$$

$Z(0), Z(1), \cdots$ are the numbers corresponding to the formulas 0, 1, $\cdots$.

20. $\mathrm{Num}(x) \equiv (En)[n \leqq x \,\&\, x = Z(n)]$.
x corresponds to one of the formulas 0, 1, $\cdots$.

21. $Z^{-1}(x) = \epsilon n[n \leqq x \,\&\, x = Z(n)]$.
If x corresponds to $\lambda fx \cdot f(\cdots n + 1 \text{ times} \cdots f(x) \cdots)$, $Z^{-1}(x)$ is n.

22. $PS(x) \equiv \mathrm{Prim}(x) \,\&\, x > 13$.
x is a *proper symbol.*

23. $PFR(x) \equiv (n)\{0 < n \leqq l(x) \rightarrow (Ev)[v \leqq x \,\&\, PS(v) \,\&\, n\, Gl\, x = R(v)] \vee (Ep, q)[0 < p, q < n \,\&\, \{n\, Gl\, x = E(p\, Gl\, x)*E(q\, Gl\, x) \vee \{n\, Gl\, x = R(1)*[p\, Gl\, x]*E(q\, Gl\, x) \,\&\, (Ev)[v \leqq x \,\&\, PS(v) \,\&\, p\, Gl\, x = R(v)]\}\}]\} \,\&\, l(x) > 0$.

x is a sequence of *formulas* of which each is either a *proper symbol* or is compounded out of the preceding ones by the operations { } ( ) and λ [ ].

24. $PF(x) \equiv (En)\{n \leqq (Pr[l(x)^2])^{x \cdot l(x)^2} \,\&\, PFR(n) \,\&\, x = [l(n)]\, Gl\, n\}$.
x is a *properly-formed formula.*[20]

25. v Geb $n, x \equiv PS(v) \,\&\, PF(x) \,\&\, (Ea, b, c)[a, b, c \leqq x \,\&\, x = a*R(1)*R(v)*E(b)*c \,\&\, PF(b) \,\&\, l(a) < n \leqq l(a) + l(b) + 4]$.
The *proper symbol* v is *bound* at the n-th point of the *properly-formed formula* x.

26. v Fr $n, x \equiv PS(v) \,\&\, PF(x) \,\&\, v = n\, Gl\, x \,\&\, n \leqq l(x) \,\&\, \overline{v \text{ Geb } n, x}$.
The *proper symbol* v occurs as a *free symbol* at the n-th point of the *properly-formed formula* x.

27. v Geb $x \equiv (En)[n \leqq l(x) \,\&\, v = n\, Gl\, x \,\&\, v$ Geb $n, x]$.
The *proper symbol* v occurs in the *properly-formed formula* x as a *bound symbol.*

28. v Fr $x \equiv (En)[n \leqq l(x) \,\&\, v$ Fr $n, x]$.
The *proper symbol* v occurs in the *properly-formed formula* x as a *free symbol.*

29. $WF(x) \equiv PF(x) \,\&\, (n)[n < l(x) \,\&\, (n+1) Gl\, x = 1 \rightarrow (Ep, q, r)\{p, q, r \leqq x \,\&\, x = p*R(1)*R[(n+2) Gl\, x]*E(q)*r \,\&\, l(p) = n \,\&\, PF(q) \,\&\, [(n+2)\, Gl\, x]$ Fr $q\}]$.
x is a *well-formed formula.*

[19] The possibility of defining number-theoretic functions by means of recursion was expounded in Skolem [1]. In that paper Skolem also showed that restricted existence and restricted generality (the restriction by an upper bound) can be expressed by recursive functions.

[20] Cf. Gödel [1] p. 183, footnote 35.

30. $x \text{ Imr } y \equiv WF(x) \mathbin{\&} \{x = y \vee (Ep, q, r, s, t)[p, q, s, t \leqq x \mathbin{\&} r \leqq y \mathbin{\&} x = p*R(1)*R(q)*E(s)*t \mathbin{\&} PS(q) \mathbin{\&} WF(s) \mathbin{\&} PS(r) \mathbin{\&} \overline{r \text{ Occ } s} \mathbin{\&} y = p*R(1)*R(r)*E(S(s, q, R(r)))*t] \vee (Ep, q, r, s, t)[p, q, r, s, t \leqq x \mathbin{\&} x = p*E(R(1)*R(q)*E(r))*E(s)*t \mathbin{\&} PS(q) \mathbin{\&} WF(r) \mathbin{\&} WF(s) \mathbin{\&} \overline{q \text{ Geb } r} \mathbin{\&} (u)[u \leqq s \mathbin{\&} u \, Fr \, s \rightarrow \overline{u \text{ Geb } r}] \mathbin{\&} y = p*S(r, q, s)*t]\}.$

$x \text{ Imc } y \equiv x \text{ Imr } y \vee y \text{ Imr } x.$

x Imr y (x Imc y) corresponds to the relation obtained from **A** conv **B** by omitting (2) and (3) (omitting (3)) in the definition of the latter.

31. $EC(x, m) = \theta(R(x), m)$ for the $\theta(z, m)$ given by Kleene [3] I when $\phi(n, x, y) = \epsilon z[z \leqq n + x \mathbin{\&} \{(x \text{ Imc } n \mathbin{\&} z = n) \vee (\overline{x \text{ Imc } n} \mathbin{\&} z = x)\}]$.

$EC(x, 0), EC(x, 1), \cdots$ is an enumeration (with repetitions) of the numbers y *convertible* into x (if x is *well-formed*).

Now let L be a given non-negative integral function of n natural numbers, and **L** a formula which λ-defines L, i.e., a formula such that, for each set $x_1, \cdots, x_n$ of natural numbers, $\mathbf{L}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ conv $\lambda fx \cdot f(\cdots m + 1 \text{ times} \cdots f(x) \cdots)$ when $m = L(x_1, \cdots, x_n)$ and (by Church-Rosser [1], Thm. 1, Cor. 2) only then. If l denotes the correspondent of **L** under our representation of well-formed formulas by natural numbers,

$$A(x_1, \cdots, x_n) = E(\cdots E(l)*E(Z(x_1)) \cdots)*E(Z(x_n))$$

is the correspondent of $\mathbf{L}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ (8, 10, 19). Hence, if $z_{x_1 \cdots x_n}$ denotes the correspondent of the formula $\lambda fx \cdot f(\cdots m + 1 \text{ times} \cdots f(x) \cdots)$, there are y's such that $EC(A(x_1, \cdots, x_n), y) = z_{x_1 \cdots x_n}$ (31). For those and only those y's Num $(EC(A(x_1, \cdots, x_n), y))$ holds (20). Hence

$$(x_1, \cdots, x_n)\ (Ey)\ \text{Num}\ (EC(A(x_1, \cdots, x_n), y))$$

and

$$Z^{-1}(EC(A(x_1, \cdots, x_n), \epsilon y[\text{Num}\ (EC(A(x_1, \cdots, x_n), y))])) = Z^{-1}(z_{x_1 \cdots x_n}) = L(x_1, \cdots, x_n) \ (21).$$

Using Kleene [3]V, the expression on the left is seen to be recursive. Thus:

(23′) *Every λ-definable non-negative integral function of natural numbers is recursive in the Herbrand-Gödel sense.*[21]

4. Recursive well-formed functions. Let L be a function of a fixed number n of natural numbers $x_1, \cdots, x_n$, of which the values $L_{x_1 \cdots x_n}$ are well-formed formulas. Let $\lambda(x_1, \cdots, x_n)$ be the function which corresponds to L under our representation of formulas by numbers, i.e., the function which the correspondent of $L_{x_1 \cdots x_n}$ is of $x_1, \cdots, x_n$. We call L recursive if $\lambda(x_1, \cdots, x_n)$

[21] This result was first announced by Church.

is recursive in the Herbrand-Gödel sense. This definition agrees with the former one, when the values of L are formulas representing natural numbers, in view of the recursiveness of Z and Z^{-1} (19, 21).

In order that L be λ-definable, it is necessary that all the values $L_{x_1 \cdots x_n}$ have the same set $\mathbf{z}_1, \cdots, \mathbf{z}_m$ of free symbols. If L is recursive, the function L' whose values are the expressions $\lambda \mathbf{z}_1 \cdots \mathbf{z}_m \cdot L_{x_1 \cdots x_n}$ (which contain no free symbols) is recursive, since

$$\lambda'(x_1, \cdots, x_n) = R(1)*R(s_1)*E(\cdots R(1)*R(s_m)*E(\lambda(x_1, \cdots, x_n)) \cdots),$$

where $s_1, \cdots, s_m$ are the numbers corresponding to the symbols $\mathbf{z}_1, \cdots, \mathbf{z}_m$, respectively. Moreover, if L' is λ-defined by $\mathbf{L}'$, then L is λ-defined by

$$\lambda \mathbf{x}_1 \cdots \mathbf{x}_n \cdot \mathbf{L}'(\mathbf{x}_1, \cdots, \mathbf{x}_n, \mathbf{z}_1, \cdots, \mathbf{z}_m).$$

These remarks reduce the problem of this section (to prove (25)) to the special case in which the values of L contain no free symbols.

In the following i and j denote the numbers corresponding to the formulas I and J, respectively:

32. $CR(x) \equiv (n)\{0 < n \leqq l(x) \rightarrow n\, Gl\, x = i \vee n\, Gl\, x = j \vee (Ep, q)[0 < p, q < n \;\&\; n\, Gl\, x = E(p\, Gl\, x)*E(q\, Gl\, x)]\} \;\&\; l(x) > 0.$

x is a sequence of *formulas* of which each is either I or J or is compounded out of the preceding ones by the operation { } ( ).

33. Comb $(x) \equiv (En)\{n \leqq (Pr[l(x)^2])^{x \cdot l(x)^2} \;\&\; CR(n) \;\&\; x = [l(n)]\, Gl\, n\}.$

x is a *combination.*

34. $C(x) = EC(x, \epsilon y\{[WF(x) \;\&\; \mathrm{Comb}\,(EC(x, y))] \vee [\overline{WF(x)} \;\&\; y = 0]\})$.[22]

If x is *well-formed,* $C(x)$ is a *combination convertible* into x.

35. $D(x) \equiv (Ep, q)[p, q \leqq x \;\&\; x = E(p)*E(q) \;\&\; WF(p) \;\&\; WF(q)].$

x corresponds to a formula of the form $\{\mathbf{P}\}$ $(\mathbf{Q})$.

36. $M_1(x) = \epsilon p[p \leqq x \;\&\; WF(p) \;\&\; (Eq)[q \leqq x \;\&\; x = E(p)*E(q)]].$

$M_2(x) = \epsilon q[q \leqq x \;\&\; WF(q) \;\&\; (Ep)[p \leqq x \;\&\; x = E(p)*E(q)]].$

If x corresponds to the formula $\{\mathbf{P}\}$ $(\mathbf{Q})$, $M_1(x)$ and $M_2(x)$ correspond to $\mathbf{P}$ and $\mathbf{Q}$, respectively.

37. $I(x) \equiv x = i.$

x corresponds to the formula I.

By the λ-definition of a relation we mean the λ-definition of the representing function of the relation (i.e., the function which is 0 or 1 according as the relation holds or not). Since a recursive relation is one of which the representing function is recursive, recursive relations among natural numbers, as well as recursive functions, are λ-definable (by (22)).

Accordingly, let $\mathbf{C}, \mathbf{D}, \mathbf{M}_1, \mathbf{M}_2, \mathbf{I}$ be formulas which λ-define $C, D, M_1, M_2, I,$

[22] By (14) and Kleene [3] V, this function is recursive, which is sufficient for our purpose. Actually, it is primitive recursive, by Gödel [1] IV, since a primitive recursive bound for y is given implicitly by the proofs of (14) and the property of $EC(x, m)$.

respectively. Using (17), choose a formula $\mathfrak{N}$ such that $\mathfrak{N}(0)$ conv $\lambda xf \cdot x(f(-1))$, $\mathfrak{N}(1)$ conv $\lambda xf \cdot x(f(-1), J)$, and a formula $\mathfrak{K}$ such that $\mathfrak{K}(0)$ conv $\lambda xf \cdot f(\mathbf{M}_1(x), f(\mathbf{M}_2(x)))$, $\mathfrak{K}(1)$ conv $\lambda xf \cdot \mathfrak{N}(\mathbf{I}(x), x, f)$; and let $\mathfrak{B} \to \lambda xf \cdot \mathfrak{K}(\mathbf{D}(x), x, f)$. By (19), there is a formula $\mathfrak{G}$ such that $\mathfrak{G}(\mathbf{x})$ conv $\mathfrak{B}(\mathbf{x}, \mathfrak{G})$ and $\mathfrak{G}(-1)$ conv I. Then $\mathfrak{G}(\mathbf{y})$ conv I if y corresponds to I, $\mathfrak{G}(\mathbf{y})$ conv J if y corresponds to J, and $\mathfrak{G}(\mathbf{y})$ conv $\mathfrak{G}(\mathbf{M}_1(\mathbf{y}), \mathfrak{G}(\mathbf{M}_2(\mathbf{y})))$ if y corresponds to a formula of the form $\{\mathbf{P}\}$ $(\mathbf{Q})$. Hence, if y corresponds to a combination $\mathbf{Y}$ whose terms are I's and J's, $\mathfrak{G}(\mathbf{y})$ conv $\mathbf{Y}$. If x corresponds to a formula $\mathbf{X}$ having no free symbols, $C(x)$ corresponds to a combination $\mathbf{Y}$ of I's and J's convertible into $\mathbf{X}$. Hence, letting $G \to \lambda x \cdot \mathfrak{G}(\mathbf{C}(x))$:

(24) *If the number x corresponds to a formula* $\mathbf{X}$ *having no free symbols,* $G(\mathbf{x})$ *conv* $\mathbf{X}$.

Now, if the function L is recursive, and if the values $L_{x_1 \cdots x_n}$ contain no free symbols, there is a formula $\mathbf{l}$ which λ-defines $\lambda(x_1, \cdots, x_n)$ (by (22)), and then $\lambda x_1 \cdots x_n \cdot G(\mathbf{l}(x_1, \cdots, x_n))$ λ-defines L. Passing to the general case by the means we have indicated:

(25) *If the function L of n natural numbers having well-formed formulas as values is recursive (i.e., if the corresponding numerical function is recursive in the Herbrand-Gödel sense), and if all the values have the same free symbols, then L is λ-definable.*

We are now in a position to infer the λ-definability of various sequences of well-formed formulas from the theory of recursive functions. We give several examples, each accompanied by a definition displaying the recursiveness of the corresponding numerical function $\lambda(x)$.[23] $a, f, \cdots$ stand for the numbers corresponding to $\mathbf{A}, \mathbf{F}, \cdots$, respectively.

(26) *The sequence* $\mathbf{A}_0, \cdots, \mathbf{A}_{k-1}, \mathbf{F}(0), \mathbf{F}(1), \cdots$ *is λ-definable (if* $\mathbf{A}_0, \cdots, \mathbf{A}_{k-1}, \mathbf{F}$ *have the same free symbols).*

$$\lambda(x) = \epsilon y[(x = 0 \;\&\; y = a_0) \vee \cdots \vee (x = k - 1 \;\&\; y = a_{k-1}) \vee (x \geqq k \;\&\; y = E(f) * E(Z(x \dot{-} k))].$$

(27) *The sequence* $\mathbf{A}_0, \cdots, \mathbf{A}_{k-1}, \mathbf{F}(0, \mathbf{A}_0, \cdots, \mathbf{A}_{k-1}), \mathbf{F}(1, \mathbf{A}_1, \cdots, \mathbf{A}_k), \cdots$*, where* $\mathbf{A}_i$ *denotes the* $(i+1)$*-th member, is λ-definable (if* $\mathbf{A}_0, \cdots, \mathbf{A}_{k-1}$ *have the same free symbols, and the free symbols of* $\mathbf{F}$ *are free symbols of* $\mathbf{A}_0$*).*

$$\lambda(0) = a_0, \cdots, \lambda(k-1) = a_{k-1}, \lambda(k + x) = E(\cdots E(E(f) * E(Z(x))) * E(\lambda(x)) \cdots) * E(\lambda(x + [k - 1])).$$

(28) *The set of formulas derivable from* $\mathbf{A}(0), \mathbf{A}(1), \cdots$ *by zero or more successive operations of passing from* $\mathbf{M}$ *and* $\mathbf{N}$ *to* $\mathbf{R}(0, \mathbf{M}, \mathbf{N}), \mathbf{R}(1, \mathbf{M}, \mathbf{N}), \cdots$ *is λ-enumerable (if the free symbols of* $\mathbf{R}$ *are free symbols of* $\mathbf{A}$*).*

[23] Here are used known recursive functions and relations, the methods of Gödel [1], Kleene [3] V, direct recursive definition by equations.

$\lambda(x) = \theta(R(E(a)*E(Z(0))), x)$, where $\theta(z, m)$ is chosen by Kleene [3] I taking

$$\phi(n, x, y) = \epsilon z\left[\left\{n|2 \;\&\; z = E\left(E\left(E(r)*E\left(Z\left(\left[\frac{n}{2}\right]\right)\right)\right)*E(x)\right)*E(y)\right\}\right.$$
$$\left.\vee\left\{n+1|2 \;\&\; z = E(a)*E\left(Z\left(\left[\frac{n+1}{2}\right]\right)\right)\right\}\right].$$

38. $EW(0) = i$ (the number corresponding to I).
$EW(x+1) = \epsilon y\{EW(x) < y \leqq Z(x) \;\&\; WF(y) \;\&\; (p)[p \leqq y \rightarrow \overline{p \text{ Fr } y}]\}$.
$EW(0)$, $EW(1)$, $\cdots$ is an enumeration of the *well-formed formulas* with no *free symbols*.

(29) *The class of well-formed formulas (having a given set of free symbols) is λ-enumerable.*

For the case of no free symbols, $\lambda(x)$ is the function $EW(x)$ which precedes; if $\mathbf{L}$ λ-enumerates the class for this case, $\lambda\mathbf{x}\cdot\mathbf{L}(\mathbf{x}, \mathbf{z}_1, \cdots, \mathbf{z}_m)$ λ-enumerates it when the set of free symbols is $\mathbf{z}_1, \cdots, \mathbf{z}_m$.

39. $NF(x) \equiv WF(x) \;\&\; (p, q, r, s, t)\{p, q, r, s, t \leqq x \;\&\; WF(q) \;\&\; WF(r) \;\&\; WF(s) \rightarrow x \neq p*E(R(1)*q*E(r))*E(s)*t\}$.
x is a *well-formed formula* in *normal form*.

40. $ENF(x)$ (defined in the same nanner as $EW(x)$ replacing $WF(x)$ by $NF(x)$).
$ENF(0)$, $ENF(1)$, $\cdots$ is an enumeration of the *well-formed formulas* with no *free symbols* in *normal form*.

41. $EN(x) = EC(ENF(1 \; Gl \; \mathrm{Dy}(x)), 2 \; Gl \; \mathrm{Dy}(x))$.
$EN(0)$, $EN(1)$, $\cdots$ is an enumeration of the *well-formed formulas* with no *free symbols* which have *normal forms*.

(30) *The class of well-formed formulas (having a given set of free symbols) which have normal forms is λ-enumerable.*[24]

This follows from 41 (or 40) in the same manner as (29) from 38.

5. **λ-definable well-formed functions.** In the extension of the notion of recursiveness to functions L of which the values are any well-formed formulas, the point of view in which interconvertible formulas are regarded as equivalent is compromised. Every well-formed formula λ-defines $2^{\aleph_0}$ functions L of n natural numbers, each corresponding to a different numerical function $\lambda(x_1, \cdots, x_n)$. Since the power of the class of recursive numerical functions is $\aleph_0$, not all functions L λ-definable by a given $\mathbf{L}$ are recursive. In order to prove a theorem like (23′), there must be added to the hypothesis of λ-definability a condition on the form of the values $L_{x_1 \cdots x_n}$ of L which selects from the formulas in which $\mathbf{L}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ is convertible that one which is $L_{x_1 \cdots x_n}$. A condition of this sort

[24] This theorem is due to Church.

which can be used here to replace the condition of representing a natural number is that of being in normal form, supplemented by a convention which removes the ambiguity in the normal form of a formula: a formula shall be in *principal normal form* if it is in normal form and the symbol following the n-th occurrence of λ is the n-th proper symbol (in the given list) which is not a free symbol of the formula.

42. $PNF(x) \equiv NF(x) \,\&\, (p, q, r)\{p, q, r \leqq x \,\&\, x = p*R(1)*R(q)*r \rightarrow$
$q = \epsilon s[s \leqq x \,\&\, PS(s) \,\&\, \overline{s \text{ Fr } x} \,\&\, \overline{s \text{ Geb } p}]\}$.
x is a *well-formed formula* in *principal normal form*.

If all the values $L_{x_1 \cdots x_n}$ are in principal normal form, and $A(x_1, \cdots, x_n)$ is chosen as in the proof of (23′), we find that $\lambda(x_1, \cdots, x_n) = EC(A(x_1, \cdots, x_n),$ $\epsilon y[PNF(EC(A(x_1, \cdots, x_n), y))])$, which is recursive in the Herbrand-Gödel sense, since $(x_1, \cdots, x_n)(Ey)PNF(EC(A(x_1, \cdots, x_n), y))$.

(31′) *Every λ-definable function of n natural numbers of which the values are well-formed formulas in principal normal form is recursive* (*i.e., the corresponding numerical function is recursive in the Herbrand-Gödel sense*).

BIBLIOGRAPHY

Alonzo Church, [1] *A set of postulates for the foundation of logic*, Ann. of Math., vol. 33 (1932), pp. 346–366. [2] A second paper under the same title, Ann. of Math., vol. 34 (1933), pp. 839–864. [3] *A proof of freedom from contradiction*, Proc. Nat. Acad. Sciences, vol. 21 (1935), pp. 275–281. [4] *An unsolvable problem of elementary number theory*, Am. Jour. Math. (to appear).

Alonzo Church and S. C. Kleene, [1] *Formal definitions in the theory of ordinal numbers*, Funda. Math. (to appear).

Alonzo Church and J. B. Rosser, [1] *Some properties of conversion*, Trans. Am. Math. Soc. (to appear).

H. B. Curry, [1] *An analysis of logical substitution*, Am. Jour. Math., vol. 51 (1929), pp. 363–384. [2] *Grundlagen der kombinatorischen Logik*, Am. Jour. Math., vol. 52 (1930), pp. 509–536, 789–834. [3] *Some additions to the theory of combinators*, Am. Jour. Math., vol. 54 (1932), pp. 551–558.

Kurt Gödel, [1] *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*, Monatsh. f. Math. u. Phys., vol. 38 (1931), pp. 173–198.

S. C. Kleene, [1] *Proof by cases in formal logic*, Ann. of Math., vol. 35 (1934), pp. 529–544. [2] *A theory of positive integers in formal logic*, Am. Jour. Math., vol. 57 (1935), pp. 153–173, 219–244. [3] *General recursive functions of natural numbers*, Math. Ann. (to appear).

S. C. Kleene and J. B. Rosser, [1] *The inconsistency of certain formal logics*, Ann. of Math., vol. 36 (1935), pp. 630–636.

J. B. Rosser, [1] *A mathematical logic without variables*, Ann. of Math., vol. 36 (1935), pp. 127–150 and this journal, vol. 1 (1935), pp. 328–355.

M. Schönfinkel, [1] *Über die Bausteine der mathematischen Logik*, Math. Ann., vol. 92 (1924), pp. 305–316.

Th. Skolem, [1] *Begründung der elementaren Arithmetik durch die rekurrierende Denkweise ohne Anwendung scheinbarer Veränderlichen mit unendlichem Ausdehnungsbereich*, Videnkapsselskapets Skrifter, 1923, I. Mat.-naturv. Kl., No. 6, pp. 1–38.

Princeton University.

SOME MORE THEOREMS CONCERNING FOURIER SERIES AND FOURIER POWER SERIES

BY G. H. HARDY AND J. E. LITTLEWOOD

1. Introduction

1.1. The principal theorem in this paper is Theorem 10: if $u(\theta)$ is periodic, with period 2π, and integrable, $s_n(x)$ is the partial sum of the Fourier series of $u(\theta)$, for $\theta = x$, $s(x)$ is arbitrary,

$$\phi(x, \theta) = \tfrac{1}{2}\{u(x + \theta) + u(x - \theta) - 2s(x)\}, \tag{1.1.1}$$

and

$$k \geqq p > 1, \tag{1.1.2}$$

then

$$\left(\sum_{1}^{\infty} \frac{|s_n(x) - s(x)|^k}{n}\right)^{1/k} \leqq K(p, k)\left(\int_0^{\pi} \frac{|\phi(x, \theta)|^p}{\theta}\, d\theta\right)^{1/p}. \tag{1.1.3}$$

This theorem is, in a sense, a theorem of 'strong summability'.[1] It is known that, if

$$\int_0^{\theta} |\phi(x, t)|^p\, dt = o(\theta), \tag{1.1.4}$$

for some $p > 1$, then

$$\sum_{1}^{n} |s_n(x) - s(x)|^k = o(n), \tag{1.1.5}$$

for every positive k. In Theorem 10 both hypothesis and conclusion are stronger. In fact (1.1.4) is equivalent to

$$\int_{\theta}^{2\theta} \frac{|\phi(x, t)|^p}{t}\, dt = o(1), \tag{1.1.6}$$

and (1.1.5) to

$$\sum_{n}^{2n} \frac{|s_\nu(x) - s(x)|^k}{\nu} = o(1); \tag{1.1.7}$$

and (1.1.6) and (1.1.7) are plainly consequences of the convergence of the integral and the series in (1.1.3).

Received November 3, 1935.

[1] See Hardy and Littlewood (**2, 4, 8**), and Zygmund (**15**), 237–241. The bold-faced numbers refer to the list of references at the end of the paper.

There are other points of difference between Theorem 10 and the theorems about strong summability. Thus (1.1.4) says the less the smaller p, and (1.1.5) says the more the larger k, in each case because of Hölder's inequality. There are no such obvious relations of inclusion between different cases of Theorem 10. The convergence of

$$\sum n^{-1} a_n^r \tag{1.1.8}$$

for given positive a_n and r, does not imply its convergence[2] for any other r; and the integral and series in (1.1.3), for different pairs of values of p and k, are similarly independent, so that no case of the theorem implies any other case in any trivial manner.

Finally, (1.1.4) is satisfied for almost all x, if $u(\theta)$ is L^p, while the integral in (1.1.3) may diverge for almost all x.

1.2. In §§2–3 we prove some 'pure inequalities' which we require later; the theorem essential for our applications is Theorem 3. In §2 we deduce this theorem from a very general theorem (Theorem 1) which we have proved elsewhere;[3] but, in view of the length and difficulty of the proof of Theorem 1, we add a direct proof of Theorem 3 in §3.

In §4 we prove our main theorem for a special class of functions, those which are boundary functions of analytic functions $f(re^{i\theta})$ regular for $r < 1$. The Fourier series of such functions are 'Fourier power series'.[4] In this case we can (as in general we cannot) include the value $p = 1$.

In §§5–6 we complete the proof of Theorem 10, for general $u(\theta)$. In §§7–9 we give a more direct proof of an analogous theorem for Fourier cosine transforms; and we conclude, in §10, with a few miscellaneous comments.

2. Inequalities

2.1. Our argument depends upon a number of special cases of a very general inequality[5] which we proved in **7** and restate here.

We suppose that

$$0 < p \leqq q, \qquad r > 0, \qquad \gamma = \alpha + \beta - 1,$$

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \qquad \alpha_0 = 1 - \frac{1}{q} + \frac{1}{r}, \qquad \beta_0 = 1 - \frac{1}{p} + \frac{1}{r},$$

$$a_0 = 0, \qquad b_0 = 0, \qquad a_n \geqq 0, \qquad b_n \geqq 0,$$

$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0,$$

[2] Thus (1.1.8) is convergent, with $a_n = (\log n + 1)^{-\alpha}$, if and only if $r > 1/\alpha$. On the other hand, if

$$a_n = n^\beta \quad (n = 2^m), \qquad a_n = 0 \quad (n \neq 2^m),$$

then (1.1.8) is convergent if and only if $r < 1/\beta$.

[3] Hardy and Littlewood (**7**), Theorem 1.

[4] For fuller explanations see Hardy and Littlewood (**5**).

[5] Hardy and Littlewood (**7**), Theorem 1.

and

$$A^p = \sum_1^\infty n^{-1}(n^\alpha a_n)^p, \qquad B^q = \sum_1^\infty n^{-1}(n^\beta b_n)^q, \qquad C^r = \sum_1^\infty n^{-1}(n^\gamma c_n)^r.$$

We allow infinite values of p, q, or r; if, for example, $p = \infty$, then A is to be interpreted as

$$\lim_{N\to\infty} \lim_{p\to\infty} \left(\sum_1^N n^{-1}(n^\alpha a_n)^p\right)^{1/p} = \max (n^\alpha a_n). \tag{2.1.1}$$

The fundamental inequality is

$$C \leqq KAB, \tag{I: 2.1.2}$$

where

$$K = K(p, q, r, \alpha, \beta). \tag{2.1.3}$$

We say that a set of conditions is necessary and sufficient for (I) if, when the conditions are satisfied, (I) is true for some K of the type (2.1.3) and all a_n, b_n, and, when they are not satisfied, (I) is false for every such K and some a_n, b_n.

In stating the theorem we distinguish between 'ordinary' and 'exceptional' cases. A case is ordinary when p, q, r are finite and $p \neq 1$, and otherwise exceptional.

THEOREM 1. (1) *It is necessary for the truth of* (I) *that*

$$p \geqq 1. \tag{2.1.4}$$

(2) [*Ordinary cases.*] *It is necessary and sufficient for the truth of* (I), *in an ordinary case, that* (2.1.4) *should be satisfied* (*so that* $p > 1$) *and also one of the four* (*mutually exclusive*) *alternative conditions*

$$\tau \leqq r \leqq p, \qquad \alpha < 1, \qquad \beta < 1; \tag{2a; 2.1.5}$$

$$p < r \leqq q, \qquad \alpha < 1, \qquad \beta \leqq \beta_0; \tag{2b; 2.1.6}$$

$$\tau \geqq 1, \qquad q < r, \qquad \alpha \leqq \alpha_0, \qquad \beta \leqq \beta_0; \tag{2c; 2.1.7}$$

$$\tau < 1, \qquad q < r \leqq \frac{\tau}{1-\tau}, \qquad \alpha \leqq \alpha_0, \qquad \beta \leqq \beta_0. \tag{2d; 2.1.8}$$

(3) [*Exceptional cases.*] *The only case*[6] *in which* $p = \infty$ *and* (I) *is true is the case*

$$p = q = r = \tau = \infty, \qquad \alpha < 1, \qquad \beta < 1.$$

When p *is finite, all exceptional cases except four are normal, in that the conditions appropriate to them can be derived from those catalogued under* (2) *by substitution of the special values of the parameters* p, q, r *and interpretation of* B *and* C, *if necessary, in accordance with the convention* (2.1.1).[7]

[6] In fact, all cases with $p = \infty$ are normal (in the sense defined below).

[7] The four cases (2a)–(2d) are mutally exclusive even when exceptional values of p, q, r are allowed, so that the definition of 'normal' is unambiguous.

The four abnormal cases are[8]

(3a) $p = q = r = 1$. *In this case the conditions are* $\alpha \leqq 1, \beta \leqq 1$ (*instead of* $\alpha < 1, \beta < 1$).

(3b) $p = 1 < r < q$. *In this case the conditions are* $\alpha < 1, \beta < \beta_0$ (*instead of* $\alpha < 1, \beta \leqq \beta_0$).

(3c) $p = 1 < r = q$. *In this case the conditions are* $\alpha \leqq 1, \beta \leqq \beta_0$ (*instead of* $\alpha < 1, \beta \leqq \beta_0$).

(3d) $p > 1, \tau > 1, r = \infty$. *In this case the conditions are* $\alpha < \alpha_0, \beta < \beta_0$ (*instead of* $\alpha \leqq \alpha_0, \beta \leqq \beta_0$).

It may be observed that $\alpha_0 < 1$ if $r > q$ and $\beta_0 < 1$ if $r > p$. Thus $\alpha < 1$ and $\beta < 1$ in all of the cases (2a)–(2d).

The case $q = \infty$

2.2. We now specialize the theorem by supposing that $q = \infty$ (so that $\tau = p$). In this case

$$B = \max (n^\beta b_n) ,$$

and $B < \infty$ means $b_n = O(n^{-\beta})$; and there is no real loss of generality in supposing that

$$b_n = n^{-\beta} = n^{\omega-1} ,$$

say.

We shall specialize a little further by supposing that

$$\alpha < 1 , \quad \beta < 1 , \quad \omega = 1 - \beta > 0 .$$

Then

$$C_n = \sum_{0<s<n} (n - s)^{\omega-1} a_s = a_n^{(\omega)} \tag{2.2.1}$$

is effectively the Riesz or Cesàro sum,[9] of order ω, formed from the series $\sum a_n$; and (I) becomes

$$(\textstyle\sum n^{-1}(n^\gamma a_n^{(\omega)})^r)^{1/r} \leqq K (\textstyle\sum n^{-1}(n^\alpha a_n)^p)^{1/p} , \tag{I$_1$: 2.2.2}$$

with

$$K = K(p, r, \alpha, \omega) . \tag{2.2.3}$$

Making these specializations in Theorem 1, and rearranging the results in a more convenient manner, we obtain

THEOREM 2. *Suppose that*

$$1 \leqq p < \infty , \quad \alpha < 1 , \quad \omega > 0 .$$

[8] The order in which these cases are catalogued is not the same as in 7; and there is no parallelism between them and (2a)–(2d).

[9] In fact C_n is Riesz's mean, with integral n, multiplied by a factor $\Gamma(\omega)$.

Riesz admits non-integral values of n, and this is important for some of the more delicate properties of his sums; but the difference is not significant here.

Then it is necessary and sufficient for the truth of (I_1), *with a* K *of type* (2.2.3), *that one or other of the sets of conditions*

$$\omega < \frac{1}{p}, \quad p \leqq r \leqq \frac{p}{1-\omega p}, \tag{2.2.4}$$

$$\omega = \frac{1}{p}, \quad p \leqq r < \infty, \tag{2.2.5}$$

$$\omega > \frac{1}{p}, \quad p \leqq r \leqq \infty \tag{2.2.6}$$

should be satisfied; except that the last $\leqq$ *in* (2.2.4) *must be changed into* $<$, *and the last* $<$ *in* (2.2.5) *into* $\leqq$, *when* $p = 1$.

Finally, if we specialize still further by supposing

$$\omega = \alpha > 0, \qquad \gamma = 0,$$

(I_1) takes the form

$$\left(\sum n^{-1}(a_n^{(\alpha)})^r\right)^{1/r} \leqq K\left(\sum n^{\alpha p-1} a_n^p\right)^{1/p} \tag{I_2: 2.2.7}$$

with

$$K = K(p, r, \alpha); \tag{2.2.8}$$

and we obtain

THEOREM 3. *If*

$$1 \leqq p < \infty, \qquad 0 < \alpha < 1,$$

then it is necessary and sufficient for the truth of (I_2) *that one or other of the sets of conditions*

$$\begin{aligned} &\alpha < \frac{1}{p}, \quad p \leqq r \leqq \frac{p}{1-\alpha p} \quad \left(\textit{and } r < \frac{1}{1-\alpha} \textit{ when } p = 1\right); \\ &\alpha = \frac{1}{p}, \quad p \leqq r < \infty; \\ &\alpha > \frac{1}{p}, \quad p \leqq r \leqq \infty \end{aligned} \tag{2.2.9}$$

should be satisfied.

3. Direct proof of Theorem 3

3.1. We have deduced Theorem 3 from Theorem 1, whose proof occupies some thirty pages of **7**; and the deduction by specialization, though straightforward, requires a good deal of attention. We therefore add a direct proof of Theorem 3, or rather of its positive clauses, which give sufficient conditions for the truth of (I_2).

We suppose first that

$$1 < p < r < \infty, \tag{3.1.1}$$

and that

$$r < \frac{p}{1 - \alpha p} \tag{3.1.2}$$

if $\alpha < 1/p$. We write[10]

$$\beta = p\alpha - 1, \qquad \gamma = \left(\frac{1}{p} - \frac{1}{r}\right)\beta, \tag{3.1.3}$$

$$S^p = \sum n^\beta a_n^p, \qquad T^r = \sum n^{-1}(a_n^{(\alpha)})^r, \tag{3.1.4}$$

so that the inequality to be proved is

$$T \leqq KS. \tag{3.1.5}$$

We can choose ρ and σ so that

$$\frac{1}{p} - 1 < \rho - \gamma < \frac{1}{p} - \alpha, \tag{3.1.6}$$

$$\frac{1}{p} - \alpha < \sigma < \frac{1}{r}.^{11} \tag{3.1.7}$$

Then

$$a_n^{(\alpha)} = \sum_{s<n} s^\gamma a_s^{1-\frac{p}{r}} \cdot s^{\rho-\gamma}(n-s)^{\sigma+\alpha-1} \cdot s^{-\rho}(n-s)^{-\sigma} a_s^{\frac{p}{r}} \leqq U^{\frac{1}{p}-\frac{1}{r}} V^{\frac{1}{p'}} W^{\frac{1}{r}}, \tag{3.1.8}$$

where p' is defined as usual by

$$p' = \frac{p}{p-1}, \qquad \frac{1}{p} + \frac{1}{p'} = 1,$$

and

$$U = \sum_{s<n} s^\beta a_s^p \leqq S^p, \tag{3.1.9}$$

$$V = \sum_{s<n} s^{(\rho-\gamma)p'} (n-s)^{(\sigma+\alpha-1)p'}, \tag{3.1.10}$$

$$W = \sum_{s<n} s^{-\rho r} (n-s)^{-\sigma r} a_s^p. \tag{3.1.11}$$

It follows from (3.1.6) and (3.1.7) that

$$(\rho - \gamma)\, p' > -1, \qquad (\sigma + \alpha - 1)\, p' > -1,$$

[10] There will be no disadvantage in using β and γ in senses different from those of §2.
[11] Observe that this choice of σ presupposes (3.1.2.).

and so

$$V \leqq Kn^{(\rho-\gamma+\sigma+\alpha-1)p'+1}, \tag{3.1.12}$$

with

$$K = K(p, r, \alpha, \rho, \sigma). \tag{3.1.13}$$

Combining (3.1.9) and (3.1.12) with (3.1.8), we obtain

$$n^{-1}(a_n^{(\alpha)})^r \leqq K\, S^{r-p}\, n^t W, \tag{3.1.14}$$

where

$$t = r\left(\rho - \gamma + \sigma + \alpha - 1 + \frac{1}{p'}\right) - 1. \tag{3.1.15}$$

Hence

$$\begin{aligned} T^r = \sum_n n^{-1}(a_n^{(\alpha)})^r &\leqq K\, S^{r-p} \sum_n n^t \sum_{s<n} s^{-\rho r}(n-s)^{-\sigma r} a_s^p \\ &= K\, S^{r-p} \sum_s s^{-\rho r} a_s^p \sum_{n>s} n^t(n-s)^{-r\sigma} \end{aligned} \tag{3.1.16}$$

But $\sigma r < 1$, by (3.1.7), and

$$t - \sigma r = r\left(\rho - \gamma + \alpha - \frac{1}{p}\right) - 1 < -1,$$

by (3.1.6), so that

$$\sum_{n>s} n^t(n-s)^{-\sigma r} < K\, s^{t-\sigma r+1}, \tag{3.1.17}$$

with a K of type (3.1.8). Finally

$$t - \sigma r - \rho r + 1 = -\gamma r + \alpha r - \frac{r}{p} = \beta,$$

by (3.1.15) and (3.1.3); so that (3.1.16) and (3.1.17) give

$$T^r \leqq K\, S^{r-p} \sum_s s^\beta a_s^p = K\, S^r.$$

Here $K = K(p, r, \alpha, \rho, \sigma)$, and $K = K(p, r, \alpha)$ when suitable values, satisfying (3.1.6) and (3.1.7), are given to ρ and σ.

3.2. Suppose next that

$$p = 1 < r < \frac{1}{1-\alpha}. \tag{3.2.1}$$

We choose ρ and σ as in §3.1, so that

$$0 < \rho - \gamma < 1 - \alpha < \sigma < \frac{1}{r},$$

and

$$s^{\rho-\gamma}(n-s)^{\sigma+\alpha-1} < n^{\rho-\gamma+\sigma+\alpha-1}.$$

Thus (3.1.8) may be replaced by

$$a_n^{(\alpha)} \leqq n^{\rho-\gamma+\sigma+\alpha-1} U^{1/r'} W^{1/r},$$

where U is defined as before (with $p = 1$), and r' like p'. The proof then proceeds as in §3.1.

There remain the marginal cases,

$$r = p; \quad p > 1, \quad \alpha < \frac{1}{p}, \quad r = \frac{p}{1-\alpha p}; \quad p > 1, \quad \alpha > \frac{1}{p}, \quad r = \infty.$$

The first of these, like the case (3.2.1) treated above, may be disposed of by an appropriate simplification of the main argument. But in this case we can go further, and find the best possible K. We therefore postpone this case, and treat it, in Theorem 4 below, as a separate theorem.

3.3. When

$$p > 1, \quad \alpha < \frac{1}{p}, \quad r = \frac{p}{1-\alpha p}, \tag{3.3.1}$$

the proof lies a little deeper.[12] It is impossible to choose σ so as to satisfy (3.1.7), and we must appeal to a theorem which Pólya[13] and we have proved elsewhere.

We write

$$b_n = n^{\alpha-1/p} a_n,$$

so that $S^p = \sum b_n^p$. Then

$$a_n^{(\alpha)} = \sum_{s<n} (n-s)^{\alpha-1} s^{\frac{1}{p}-\alpha} b_s \leqq n^{\frac{1}{p}-\alpha} \sum_{s<n} (n-s)^{\alpha-1} b_s = n^{\frac{1}{p}-\alpha} b_n^{(\alpha)}.$$

Hence

$$T^{\frac{p}{1-\alpha p}} = \sum n^{-1} (a_n^{(\alpha)})^{\frac{p}{1-\alpha p}} \leqq \sum (b_n^{(\alpha)})^{\frac{p}{1-\alpha p}} \leqq K\left(\sum b_n^p\right)^{\frac{1}{1-\alpha p}} = KS^{\frac{p}{1-\alpha p}},$$

by the theorem referred to.

3.4. There remains the case

$$p > 1, \quad \alpha > \frac{1}{p}, \quad r = \infty. \tag{3.4.1}$$

The theorem then asserts that

$$a_n^{(\alpha)} = \sum_{s<n} (n-s)^{\alpha-1} a_s \leqq KS. \tag{3.4.2}$$

[12] We are in case V of Theorem 1: see **7**, §§7–8.

[13] Hardy, Littlewood and Pólya (**9**, Theorem 5).

Now

$$(3.4.3)\qquad \sum_{s<n}(n-s)^{\alpha-1}a_s=\sum_{s<n}s^{\alpha-\frac{1}{p}}a_s\cdot s^{\frac{1}{p}-\alpha}(n-s)^{\alpha-1}$$
$$\leqq S\left(\sum_{s<n}s^{\frac{p'}{p}-\alpha p'}(n-s)^{p'(\alpha-1)}\right)^{\frac{1}{p'}}.$$

Since

$$\frac{p'}{p}-\alpha p'=p'-1-\alpha p'=-1+p'(1-\alpha)>-1,$$
$$p'(\alpha-1)=p'\left(\alpha-\frac{1}{p}\right)-1>-1,$$

and the sum of these indices is -1, the second factor on the right hand side of (3.4.3) is bounded; and this proves (3.4.2).

3.5. When $r=p$, we can prove the more precise result which follows.

THEOREM 4. *If* $p\geqq 1$, $0<\alpha<1$, *then*

$$(3.5.1)\qquad \sum n^{-1}(a_n^{(\alpha)})^p\leqq(\pi\operatorname{cosec}\alpha\pi)^p\sum n^{\alpha p-1}a_n^p.$$

The constant factor is the best possible.

(1) If $p>1$, we take

$$(3.5.2)\qquad \rho=-\frac{\alpha}{p'},\qquad \sigma=\frac{1-\alpha}{p}.$$

These values satisfy (3.1.6) and (3.1.7).[14] Then

$$a_n^{(\alpha)}\leqq\left(\sum_{s<n}s^{\rho p'}(n-s)^{(\sigma+\alpha-1)p'}\right)^{1/p'}\left(\sum_{s<n}s^{-\rho p}(n-s)^{-\sigma p}a_s^p\right)^{1/p}=V^{1/p'}W^{1/p}.$$

Here

$$V=\sum_{s<n}s^{-\alpha}(n-s)^{\alpha-1}<\int_0^n x^{-\alpha}(n-x)^{\alpha-1}\,dx=\pi\operatorname{cosec}\alpha\pi.$$

Hence

$$(3.5.3)\qquad T^p=\sum n^{-1}(a_n^{(\alpha)})^p\leqq(\pi\operatorname{cosec}\alpha\pi)^{p-1}\sum_n n^{-1}\sum_{s<n}s^{-\rho p}(n-s)^{-\sigma p}a_s^p$$
$$=(\pi\operatorname{cosec}\alpha\pi)^{p-1}\sum_s s^{-\rho p}a_s^p\sum_{n>s}n^{-1}(n-s)^{-\sigma p}.$$

But

$$(3.5.4)\qquad \sum_{n>s}n^{-1}(n-s)^{-\sigma p}=\sum_{n>s}n^{-1}(n-s)^{\alpha-1}<\int_s^\infty x^{-1}(x-s)^{\alpha-1}\,dx$$
$$=\pi\operatorname{cosec}\alpha\pi\cdot s^{\alpha-1}$$

[14] γ is now 0.

and

$$-\rho p + \alpha - 1 = -(p-1)\alpha + \alpha - 1 = \alpha p - 1. \tag{3.5.5}$$

Hence (3.5.1) follows from (3.5.3).

If $p = 1$, $r = 1$, the equations (3.5.2) reduce to $\rho = 0$, $\sigma = 1 - \alpha$, and the conditions (3.1.6) and (3.1.7) are not satisfied. But in this case

$$\sum_n n^{-1} a_n^{(\alpha)} = \sum_n n^{-1} \sum_{s<n} (n-s)^{\alpha-1} a_s = \sum_s a_s \sum_{n>s} n^{-1}(n-s)^{\alpha-1}$$
$$\leqq \pi \operatorname{cosec} \alpha\pi \sum_s s^{\alpha-1} a_s,$$

and (3.5.1) is still correct.

We have now completed the proof of the main clause of Theorem 4 (and so that of Theorem 3). To prove the constant in Theorem 4 the best possible, we take $a_n = n^{-\alpha-\delta}$, where δ is small and positive. Then

$$a_n^{(\alpha)} = \sum_{s<n} (n-s)^{\alpha-1} s^{-\alpha-\delta} \sim \frac{\Gamma(\alpha)\Gamma(1-\alpha-\delta)}{\Gamma(1-\delta)} n^{-\delta};$$

and it follows, by an argument of a familiar type,[15] that the constant cannot be less than

$$\lim_{\delta \to 0} \left(\frac{\Gamma(\alpha)\Gamma(1-\alpha-\delta)}{\Gamma(1-\delta)} \right)^p = (\pi \operatorname{cosec} \alpha\pi)^p.$$

The Riesz mean of a_n is not $a_n^{(\alpha)}$ but $a_n^{(\alpha)}/\Gamma(\alpha)$. If $a_n^{(\alpha)}$ were actually the Riesz mean, the constant would be $(\Gamma(1-\alpha))^p$.

3.6. All these theorems have naturally their analogues for integrals. In particular we require

THEOREM 5. *Suppose that p, r, and α satisfy the conditions of Theorem 3; that $f(x) \geqq 0$; and that*

$$f_\alpha(x) = \frac{1}{\Gamma(\alpha)} \int_0^x (x-y)^{\alpha-1} f(y)\, dy \tag{3.6.1}$$

is the Liouville integral of $f(x)$ of order α, with origin 0. Then

$$\left(\int_0^\infty x^{-1} (f_\alpha(x))^r\, dx \right)^{1/r} \leqq K \left(\int_0^\infty x^{\alpha p - 1} (f(x))^p\, dx \right)^{1/p}. \tag{3.6.2}$$

When $r = p$ we can take $K = \Gamma(1-\alpha)$, and this is then the best possible value of K.

The proof is the same except for trivial simplifications.

Finally, taking

$$p > 1, \quad \alpha = 1/p; \qquad p > 1, \quad \alpha = 1/p',$$

in Theorems 3 and 5, we obtain two theorems which are particularly important for our applications.

[15] See Hardy, Littlewood and Pólya (**10**), 232.

THEOREM 6. *If*

$$p > 1, \qquad p \leqq r < \infty,$$

and $a_n \geqq 0, f(x) \geqq 0$, *then*

$$(\sum n^{-1}(a_n^{(1/p)})^r)^{1/r} \leqq K(\sum a_n^p)^{1/p}, \tag{3.6.3}$$

$$\left(\int_0^\infty x^{-1}(f_{1/p}(x))^r\, dx\right)^{1/r} \leqq K\left(\int_0^\infty (f(x))^p\, dx\right)^{1/p}. \tag{3.6.4}$$

THEOREM 7. *If*

$$1 < p < 2, \qquad p \leqq r \leqq \frac{p}{2-p},$$

or

$$p = 2, \qquad 2 \leqq r < \infty,$$

or

$$p > 2, \qquad p \leqq r \leqq \infty,$$

and $a_n \geqq 0, f(x) \geqq 0$, *then*

$$(\sum n^{-1}(a_n^{(1/p')})^r)^{1/r} \leqq K(\sum n^{p-2}a_n^p)^{1/p}, \tag{3.6.5}$$

$$\left(\int_0^\infty x^{-1}(f_{1/p'}(x))^r\, dx\right)^{1/r} \leqq K\left(\int_0^\infty x^{p-2}(f(x))^p\, dx\right)^{1/p}. \tag{3.6.6}$$

In each theorem $K = K(p, r)$.

4. Theorems on power series

4.1. We suppose now that

$$f(z) = \sum_0^\infty c_n z^n \tag{4.1.1}$$

is an analytic function regular for $r = |z| < 1$, that $p \geqq 1$, and that

$$\int_{-\pi}^{\pi} |f(z)|^p\, |1 - z|^{p-1}\, d\theta = \int_{-\pi}^{\pi} |f(re^{i\theta})|^p\, |1 - re^{i\theta}|^{p-1}\, d\theta \tag{4.1.2}$$

is bounded for $r < 1$.

We may always suppose, if we please, that $c_0 = 0$, since the theorems which we prove under this restriction may be extended to the general case by considering $z\,f(z)$ instead of $f(z)$. Series in which n occurs as a denominator are extended over the range 1 to ∞.

If $p = 1$, $f(z)$ belongs to the (complex) class L, and all the standard relations hold between the function, its boundary values, and its coefficients. In particular,

$$|c_n| \leqq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|\, d\theta$$

and

$$\sum \frac{|c_n|}{n} \leqq \frac{A}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|\, d\theta ,$$

where A is a constant.[16] Hence

$$\left(\sum \frac{|c_n|^k}{n}\right)^{1/k} \leqq \frac{A}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|\, d\theta$$

for $1 \leqq k \leqq \infty$.

The situation is not quite so simple when $p > 1$, since $f(z)$ does not usually belong to L. If however $0 < \lambda < 1$, then[17]

$$\int |f|^\lambda\, d\theta = \int |f|^\lambda |1 - z|^{\frac{(p-1)\lambda}{p}} \cdot |1 - z|^{-\frac{(p-1)\lambda}{p}}\, d\theta$$

$$\leqq \left(\int |f|^p |1 - z|^{p-1}\, d\theta\right)^{\frac{\lambda}{p}} \left(\int |1 - z|^{-\frac{(p-1)\lambda}{p-\lambda}}\, d\theta\right)^{\frac{p-\lambda}{p}},$$

and $(p - 1)\lambda < p - \lambda$, so that the second factor is bounded. Hence $f(z)$ belongs to the class L^λ, and has a boundary function

$$F(\theta) = f(e^{i\theta})$$

of L^λ. Also, since $(1 - z)^{1/p'} f(z)$ is a power series of the class L^p, and $(1 - e^{i\theta})^{1/p'} F(\theta)$ is its boundary function, we have

$$\int |F(\theta)|^p |1 - e^{i\theta}|^{p-1}\, d\theta = \lim_{r \to 1} \int |f(z)|^p |1 - z|^{p-1}\, d\theta < \infty .$$

On the other hand, if

$$\int |F(\theta)|^p |1 - e^{i\theta}|^{p-1}\, d\theta < \infty,$$

then $(1 - e^{i\theta})^{1/p'} F(\theta)$ is L^p, and is the boundary function of a function $(1 - z)^{1/p'} g(z)$ of the complex class L^p; and $g(z)$ must be $f(z)$, since $F(\theta)$ is the boundary function of $f(z)$. Hence

$$\int |f(z)|^p |1 - z|^{p-1}\, d\theta$$

is bounded. Finally, since the ratio

$$\int |F(\theta)|^p |\theta|^{p-1}\, d\theta \quad : \quad \int |F(\theta)|^p |1 - e^{i\theta}|^{p-1}\, d\theta$$

[16] Hardy and Littlewood (**3**), 208.

[17] The range of θ is always supposed to be $(-\pi, \pi)$.

lies between positive bounds depending only on p, our condition on $f(z)$ is equivalent to the condition

$$\int | F(\theta) |^p | \theta |^{p-1} d\theta < \infty . \tag{4.1.3}$$

4.2. THEOREM 8. *If*

$$1 \leqq p \leqq k < \infty$$

and the integral (4.1.2) *is bounded, or* $F(\theta)$ *satisfies* (4.1.3), *then*

$$\left(\sum \frac{| c_n |^k}{n}\right)^{1/k} \leqq K\left(\int | F(\theta) |^p | \theta |^{p-1} d\theta\right)^{1/p}, \tag{4.2.1}$$

with $K = K(p, k)$.

(1) We have already disposed of the case $p = 1$; in this case we may include the value $k = \infty$.

(2) We suppose then that $p > 1$. We shall use (besides the theorems of §3) three known theorems concerning Fourier series, expressed by the inequalities

$$\left(\sum | u_n |^{p'}\right)^{1/p'} \leqq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} | h |^p d\theta\right)^{1/p} \qquad (1 < p \leqq 2), \tag{4.2.2}$$

$$\sum | n |^{p-2} | u_n |^p \leqq K \int | h |^p d\theta \qquad (1 < p \leqq 2), \tag{4.2.3}$$

$$\sum | u_n |^p \leqq K \int | h |^p | \theta |^{p-2} d\theta \qquad (p \geqq 2). \tag{4.2.4}$$

In these theorems, of which the first is due to Hausdorff[18] and the second and third to ourselves,[19]

$$\sum_{-\infty}^{\infty} u_n e^{ni\theta}$$

is the complex Fourier series of a function $h(\theta)$ for which the integral on the right hand side is finite.

We must distinguish the cases $p \leqq 2$ and $p \geqq 2$.

(3) Suppose first that $1 < p \leqq 2$. It is plain from Hölder's inequality that, if (4.2.1) is true for $k = k_1$ and for $k = k_2$ (and the same p), it is true for $k_1 \leqq k \leqq k_2$. It is therefore sufficient to prove it (a) when

$$p \leqq k \leqq \frac{p}{2 - p} \tag{4.2.5}$$

and (b) when[20]

$$k \geqq p' . \tag{4.2.6}$$

[18] Hausdorff (**11**); Zygmund (**15**), 189–192, 200–202.

[19] Hardy and Littlewood (**3**), Theorems 5 and 3; Zygmund (**15**), 202–215.

[20] The two ranges overlap or abut when $p \geqq \frac{3}{2}$, and then no appeal to Hölder's inequality is necessary.

The function

$$\phi(z) = \sum n^{-1/p} z^n$$

is regular for $r \leqq 1$, $z \neq 1$, and has no zeros, except $z = 0$, in or on the unit circle.[21] Near[22] $z = 1$

$$\phi(z) \sim \frac{\Gamma(1/p')}{(1 - z)^{1/p'}},$$

and the ratio

$$| \phi(e^{i\theta}) | : | \theta |^{-1/p'}$$

lies between positive bounds $K(p)$.

If

$$g(z) = \sum b_n z^n = \frac{f(z)}{\phi(z)}, \qquad f(z) = \phi(z)\, g(z),$$

then

$$c_n = \sum_{s<n} (n - s)^{-1/p}\, b_s = b_n^{(1/p')},$$

in the notation of §2.2. The function $g(z)$ is regular and belongs to L^p, and has a boundary function $g(e^{i\theta}) = G(\theta)$; and the ratio

$$| G(\theta) | : | \theta |^{1/p'} | F(\theta) |$$

lies (for almost all θ) between positive bounds $K(p)$.

We now distinguish cases (*a*) and (*b*). In case (*a*) we use Theorem 7 and the second of the three theorems quoted in (2), viz. (4.2.3). These give

$$\left(\sum \frac{| c_n |^k}{n}\right)^{1/k} = \left(\sum \frac{| b_n^{(1/p')} |^k}{n}\right)^{1/k} \leqq K\left(\sum n^{p-2} | b_n |^p\right)^{1/p}$$
$$\leqq K\left(\int | G |^p \, d\theta\right)^{1/p} \leqq K\left(\int | F |^p | \theta |^{p-1} \, d\theta\right)^{1/p}.$$

In case (*b*) we use Theorem 6 (with p' in place of p) and the first of the theorems of (2), viz. (4.2.2). We thus obtain

$$\left(\sum \frac{| c_n |^k}{n}\right)^{1/k} = \left(\sum \frac{| b_n^{(1/p')} |^k}{n}\right)^{1/k} \leqq K\left(\sum | b_n |^{p'}\right)^{1/p'}$$
$$\leqq K\left(\int | G |^p \, d\theta\right)^{1/p} \leqq K\left(\int | F |^p | \theta |^{p-1} \, d\theta\right)^{1/p}.$$

[21] This is a case of 'Kakeya's Theorem'.

[22] In fact

$$\phi(z) - \Gamma\left(\frac{1}{p'}\right)\left(\log \frac{1}{z}\right)^{-1/p'}$$

is regular for $z = 1$. See for example Lindelöf (**12**), 138.

Thus the theorem is proved in cases (*a*) and (*b*), and so whenever $p \leqq 2$.

(4) When $p \geqq 2$ we use

$$\psi(z) = \sum n^{-1/p'} z^n$$

instead of $\phi(z)$. If

$$f = \psi g, \qquad g = \sum b_n z^n,$$

then

$$c_n = b_n^{(1/p)}.$$

The function $g(z)$ has a boundary function $g(e^{i\theta}) = G(\theta)$, and the ratio

$$| G(\theta) | : | \theta |^{1/p} | F(\theta) |$$

lies between positive bounds $K(p)$. Hence, using now Theorem 6 and the third of the theorems of (2), viz. (4.2.4), we obtain

$$\left(\sum \frac{| c_n |^k}{n}\right)^{1/k} = \left(\sum \frac{| b_n^{(1/p)} |^k}{n}\right)^{1/k} \leqq \left(\sum | b_n |^p\right)^{1/p}$$

$$\leqq K\left(\int | G |^p | \theta |^{p-2} \, d\theta\right)^{1/p} \leqq K\left(\int | F |^p | \theta |^{p-1} \, d\theta\right)^{1/p},$$

thus completing the proof of Theorem 8.

The result is not true, for any $p > 1$, when $k = \infty$. It would imply that $b_n^{(1/p')}$ is bounded for any $g(z) = \sum b_n z^n$ of the class L^p, and it is not difficult to construct an example to the contrary.

4.3. Theorem 9. *Suppose that $F(\theta)$ is the boundary function of an analytic function $f(z) = \sum c_n z^n$ of the class L; that*

$$s_n(x) = \sum_0^n c_\nu e^{\nu i x};$$

that $k \geqq p \geqq 1$; and that

$$| \theta |^{-1} | F(x + \theta) - s(x) |^p$$

is, for a given x and $s(x)$, integrable in θ. Then

$$\left(\sum_1^\infty \frac{| s_n(x) - s(x) |^k}{n}\right)^{1/k} \leqq K\left(\int_{-\pi}^{\pi} \frac{| F(x + \theta) - s(x) |^p}{| \theta |} \, d\theta\right)^{1/p},$$

with $K = K(p, k)$.

We may suppose $x = 0$. We have then only to write

$$f(z) - s(0) = h(z),$$

$$g(z) = \frac{h(z)}{1 - z} = \sum (s_n(0) - s(0)) z^n$$

and to apply Theorem 8 to $g(z)$.

5. Extension to general Fourier series

5.1. It is natural to expect that, when $p > 1$, there will be a theorem for general Fourier series corresponding to Theorem 9.

Let us suppose that $p > 1$; that $u(\theta)$ is a periodic function of θ of the class L^p; that

$$u(\theta) \sim \sum_{-\infty}^{\infty} c_n\, e^{nix}$$

(or

$$u(\theta) \sim \tfrac{1}{2}\, a_0 + \sum_{1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta))$$

is the Fourier series of $u(\theta)$; that

$$s_n(x) = \sum_{-n}^{n} c_\nu\, e^{\nu ix}$$

(or

$$s_n(x) = \tfrac{1}{2}\, a_0 + \sum_{1}^{n} (a_\nu \cos \nu x + b_\nu \sin \nu x)),$$

and that $\phi(x, \theta)$ is defined as in (1.1.1). We shall prove

THEOREM 10. *If $k \geqq p > 1$ and $|\,\theta\,|^{-1}\,|\,\phi(x, \theta)\,|^p$ is integrable in θ, for a given x and $s(x)$, then*

$$\left(\sum_{1}^{\infty} \frac{|\,s_n(x) - s(x)\,|^k}{n}\right)^{1/k} \leqq K\left(\int_0^\pi \frac{|\,\phi(x, \theta)\,|^p}{\theta}\, d\theta\right)^{1/p}, \tag{5.1.1}$$

with $K = K(p, k)$.

5.2. We may make the usual formal simplifications, supposing $x = 0$, $s(x) = 0$, and $u(\theta)$ real and even, so that

$$u(\theta) \sim \tfrac{1}{2}a_0 + \sum a_n \cos n\theta,$$

$$\phi(x, \theta) = u(\theta),$$

$$s_n = s_n(x) = s_n(0) = \tfrac{1}{2}a_0 + \sum_{1}^{n} a_\nu .$$

We have then to prove that

$$\left(\sum_{1}^{\infty} \frac{|\,s_n\,|^k}{n}\right)^{1/k} \leqq K\left(\int_0^\pi \frac{|\,u(\theta)\,|^p}{\theta}\, d\theta\right)^{1/p}. \tag{5.2.1}$$

The function u has a conjugate v, odd and of L^p, defined by

$$v(\theta) = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \cot \frac{1}{2}\,(\phi - \theta)\, u(\phi)\, d\phi . \tag{5.2.2}$$

The associated harmonic function vanishes at the origin.

Let us assume for a moment that we have proved that

$$\int_0^\pi \frac{|v(\theta)|^p}{\theta}\,d\theta \leqq K(p)\int_0^\pi \frac{|u(\theta)|^p}{\theta}\,d\theta \tag{5.2.3}$$

whenever the integral on the right is finite. Then $F(\theta) = u(\theta) + iv(\theta)$ is the boundary function of an analytic function $f(z) = \sum c_n z^n$ satisfying the conditions of Theorem 9, and $s_n = c_0 + c_1 + \cdots + c_n$. Hence

$$\left(\sum \frac{|s_n|^k}{n}\right)^{1/k} \leqq K\left(\int \frac{|F(\theta)|^p}{|\theta|}\,d\theta\right)^{1/p} \leqq K\left(\int \frac{|u(\theta)|^p}{|\theta|}\,d\theta\right)^{1/p},$$

with $K = K(p, k)$.

The proof of Theorem 10 is thus reduced to the proof of (5.2.3).

6. Theorems on conjugate functions

6.1. It is well known[23] that a function $U(x)$ of $L^p(-\infty, \infty)$, where $p > 1$, possesses a conjugate $V(x)$, defined for almost all x by

$$V(x) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{U(y)}{y - x}\,dy\,, \tag{6.1.1}$$

which also belongs to $L^p(-\infty, \infty)$. The integral is a Lebesgue integral at infinity and a principal value at $y = x$.

When $U(x)$ is *even* (a hypothesis essential in the sequel), we may also write

$$V(x) = -\frac{1}{\pi}\int_0^{\infty} \frac{2x}{y^2 - x^2}\,U(y)\,dy\,. \tag{6.1.2}$$

This integral exists under wider conditions than that in (6.1.1). Suppose, for example, that $U(x)$ is L^p in every finite positive interval, and that $x^\alpha U$, where

$$\alpha > -1 - \frac{1}{p},$$

is $L^p(0, \infty)$. Then the integral converges as a principal value (for almost all x) across $y = |x|$; and, since

$$\int^\infty \frac{|U|}{y^2}\,dy \leqq \left(\int^\infty (y^\alpha |U|)^p\,dy\right)^{1/p}\left(\int^\infty y^{-(2+\alpha)p'}\,dy\right)^{1/p'}$$

and

$$(2 + \alpha)p' > \left(1 - \frac{1}{p}\right)p' = 1\,,$$

it converges absolutely at infinity. We may therefore define $V(x)$ by (6.1.2).

THEOREM 11. *If*

$$-1 - \frac{1}{p} < \alpha < \frac{1}{p'} = 1 - \frac{1}{p}, \tag{6.1.3}$$

[23] M. Riesz (**13**); Zygmund (**15**), 147–149.

$x^\alpha U(x)$ is $L^p(0, \infty)$, and $V(x)$ is defined by (6.1.2), *then $x^\alpha V(x)$ is $L^p(0, \infty)$, and*

$$\int_0^\infty (x^\alpha | V |)^p \, dx \leqq K \int_0^\infty (x^\alpha | U |)^p \, dx, \tag{6.1.4}$$

with $K = K(p, \alpha)$.

We denote by $V^*(x)$ the conjugate of the even function $| x |^\alpha U(x)$. Then $V^*(x)$ is L^p and

$$\int_0^\infty | V^* |^p \, dx \leqq K \int_0^\infty (x^\alpha | U |)^p \, dx .$$

It is therefore sufficient to prove that

$$\int_0^\infty | V^* - x^\alpha V |^p \, dx \leqq K \int_0^\infty (x^\alpha | U |)^p \, dx . \tag{6.1.5}$$

Now, when $x > 0$,

$$V^* - x^\alpha V = \frac{2}{\pi} \int_0^\infty M(y, x) y^\alpha U(y) \, dy ,$$

where

$$M(y, x) = \frac{x}{y^\alpha} \frac{y^\alpha - x^\alpha}{y^2 - x^2}.$$

This function has a fixed sign, viz. that of α, and is homogeneous of degree -1; and

$$\int_0^\infty | M(y, 1) | y^{-1/p} \, dy = \int_0^\infty \left| \frac{y^\alpha - 1}{y^2 - 1} \right| y^{-\alpha - 1/p} \, dy < \infty$$

when α satisfies (6.1.3). Hence[24] (6.1.5) is true under the conditions of the theorem.

In particular, when $\alpha = -1/p$, we obtain

$$\int_0^\infty \frac{| V |^p}{x} \, dx \leqq K(p) \int_0^\infty \frac{| U |^p}{x} \, dx . \tag{6.1.5}$$

6.2. Theorem 12. *If $p > 1$, $u(\theta)$ is periodic and even, $\theta^{-1} | u(\theta) |^p$ is integrable, and $v(\theta)$ is defined by* (5.2.2), *then*

$$\int_0^\pi \frac{| v(\theta) |^p}{\theta} \, d\theta \leqq K(p) \int_0^\pi \frac{| u(\theta) |^p}{\theta} \, d\theta . \tag{6.2.1}$$

We have

$$\begin{aligned} v(\theta) &= - \frac{1}{2\pi} \int_0^\pi \left(\cot \frac{1}{2} (\phi - \theta) - \cot \frac{1}{2} (\phi + \theta) \right) u(\phi) \, d\phi \\ &= \frac{1}{\pi} \int_0^\pi \frac{\sin \theta}{\cos \phi - \cos \theta} u(\phi) \, d\phi . \end{aligned} \tag{6.2.2}$$

[24] Hardy, Littlewood, and Pólya (**10**), Theorem 319.

If we write

$$x = \tan \tfrac{1}{2}\theta, \qquad y = \tan\tfrac{1}{2}\phi, \qquad u(\phi) = U(y), \qquad v(\theta) = V(x),$$

then (6.2.2) becomes (6.1.2). Also

$$\int_0^\pi \frac{|u(\theta)|^p}{\sin\theta}\,d\theta = \int_0^\infty \frac{|U(x)|^p}{x}\,dx, \qquad \int_0^\pi \frac{|v(\theta)|^p}{\sin\theta}\,d\theta = \int_0^\infty \frac{|V(x)|^p}{x}\,dx,$$

and therefore, by Theorem 11,

$$\int_0^\pi \frac{|v(\theta)|^p}{\sin\theta}\,d\theta \leqq K(p) \int_0^\pi \frac{|u(\theta)|^p}{\sin\theta}\,d\theta.$$

This implies (6.2.1).

In proving this theorem, we have completed the proof of Theorem 10.

6.3. It is to be observed that the truth of Theorems 11 and 12 depends essentially on the hypothesis that $U(x)$ and $u(\theta)$ are even. It is not true, without this restriction, that the integrability of $\theta^{-1} \mid u(\theta) \mid^p$ involves that of $\theta^{-1} \mid v(\theta) \mid^p$. Suppose for example that

$$u(\theta) = \sum \frac{\sin n\theta}{n\,(\log n)^\beta}, \qquad v(\theta) = -\sum \frac{\cos n\theta}{n\,(\log n)^\beta},$$

where β is positive. Then $u(\theta)$ behaves like a multiple of $\mid \log \theta \mid^{-\beta}$ for small positive θ, and $\mid \theta \mid^{-1} \mid u \mid^p$ is integrable if (and only if) $\beta p > 1$. On the other hand $v(\theta)$ behaves like a multiple of $\mid \log \theta \mid^{1-\beta}$, if $\beta < 1$, and $v(\theta) - v(0)$ behaves in this way if $\beta > 1$; and neither $\mid \theta \mid^{-1} \mid v \mid^p$ nor $\mid \theta \mid^{-1} \mid v(\theta) - v(0) \mid^p$ is integrable unless $\beta p > p + 1$.

We can show, by an argument like that of §6.1, but based upon the formula (6.1.1) instead of upon (6.1.2), that the conclusion of Theorem 11 is true, for general $U(x)$, when $-1/p < \alpha < 1/p'$; but the value $-1/p$ of α, the critical value for our purpose, is excluded.

7. Fourier transforms

7.1. Our proof of Theorem 10 is comparatively simple (granted the inequalities of §§2–3) but very indirect, and it is natural to ask for a proof independent of the theory of analytic functions. For the sake of variety we give here not this proof but the proof of the analogous theorem for Fourier cosine transforms. To simplify the formulae, we suppose throughout that $k = p$.

We use the notion of a 'limit in mean' or 'strong limit', with index p, of a function $s_a(x)$. This, if it exists, is a function $s(x)$ such that

$$\lim_{a\to\infty} \int_0^X \mid s_a(x) - s(x) \mid^p dx = 0$$

for every positive and finite X. We write, after Wiener,

$$s(x) = \operatorname*{l.i.m.}_{a\to\infty}\, s_a(x).$$

A limit in mean is, apart from null sets, unique.

THEOREM 13. *Suppose that* $p > 1$,

$$\int_0^\infty \frac{|f(x)|^p}{x}\,dx < \infty, \tag{7.1.1}$$

and

$$s_a(x) = \int_0^a f(t)\,\frac{\sin xt}{t}\,dt. \tag{7.1.2}$$

Then $s_a(x)$ *has a limit in mean* $s(x)$ *when* $a \to \infty$, *and*

$$\int_0^\infty \frac{|s(x)|^p}{x}\,dx \leqq K(p)\int_0^\infty \frac{|f(x)|^p}{x}\,dx. \tag{7.1.3}$$

7.2. There is a simple proof, which we have not succeeded in generalizing, in the case $p = 2$.

Suppose first that $f(x) = 0$ for $x > c$. Then

$$s_a(x) = \int_0^c f(t)\,\frac{\sin xt}{t}\,dt = \int_0^\infty f(t)\,\frac{\sin xt}{t}\,dt = s(x)$$

for $a > c$, so that $s(x)$ is the limit of $s_a(x)$ in the ordinary sense. Also,

$$\begin{aligned}\int_0^\xi \frac{(s(x))^2}{x}\,dx &= \int_0^\xi \frac{dx}{x}\int_0^c f(t)\,\frac{\sin xt}{t}\,dt\int_0^c f(u)\,\frac{\sin xu}{u}\,du\\ &= \int_0^\infty\int_0^\infty \frac{f(t)f(u)}{tu}\,dt\,du\int_0^\xi \frac{\sin xt\sin xu}{x}\,dx.\end{aligned} \tag{7.2.1}$$

The inner integral is

$$\begin{aligned}\frac{1}{2}\int_0^\xi \frac{1-\cos(t+u)x}{x}\,dx - \frac{1}{2}\int_0^\xi \frac{1-\cos|t-u|\,x}{x}\,dx\\ = \frac{1}{2}\int_{|t-u|\xi}^{(t+u)\xi}\frac{1-\cos w}{w}\,dw,\end{aligned}$$

and is positive and less than

$$\log\frac{t+u}{|t-u|}.$$

Hence, if we write $g(t) = t^{-\frac{1}{2}}f(t)$, so that $g(t)$ is L^2, we have

$$\int_0^\xi \frac{(s(x))^2}{x}\,dx \leqq \int_0^\infty\int_0^\infty M(t,u)\,g(t)\,g(u)\,dt\,du, \tag{7.2.2}$$

where

$$M(t,u) = \frac{1}{\sqrt{(tu)}}\log\frac{t+u}{|t-u|}. \tag{7.2.3}$$

Finally, since M is homogeneous of degree -1, and

$$m = \int_0^\infty M(t, 1)t^{-\frac{1}{2}}\,dt = \int_0^\infty \frac{1}{t}\log\frac{t+1}{|t-1|}\,dt = \frac{1}{2}\pi^2 < \infty,$$

(7.2.2) implies[25]

$$\int_0^\infty \frac{(s(x))^2}{x}\,dx \leqq m\int_0^\infty (g(t))^2\,dt = m\int_0^\infty \frac{(f(t))^2}{t}\,dt. \tag{7.2.4}$$

Passing to the general case, we observe that

$$s_b(x) - s_a(x) = \int_a^b f(t)\,\frac{\sin xt}{t}\,dt,$$

and so, after (7.2.4),

$$\int_0^\infty \frac{(s_b(x) - s_a(x))^2}{x}\,dx \leqq m\int_a^b \frac{(f(t))^2}{t}\,dt,$$

which tends to 0 when a and b tend to infinity. À fortiori

$$\int_0^X (s_b(x) - s_a(x))^2\,dx \to 0$$

if $0 < X < \infty$. It now follows in the usual manner that $s(x)$ exists for almost all x, and that

$$\int_0^\infty \frac{(s(x))^2}{x}\,dx = \lim_{a\to\infty}\int_0^\infty \frac{(s_a(x))^2}{x}\,dx \leqq m\int_0^\infty \frac{(f(t))^2}{t}\,dt.$$

We observe here, in order to avoid repetition, that the last stage of the argument would run quite similarly for general p. When we have proved the analogue of (7.2.4), with general p and $f = 0$ for $t > c$, the rest of the theorem will follow.

8. Lemmas for the proof of Theorem 13

8.1. Lemma α. *If $f(x) \geqq 0$, $p > 1$, $r > 1$,*

$$f_1(x) = \int_0^x f(t)\,dt, \qquad f_2(x) = \int_x^\infty \frac{f(t)}{t}\,dt,$$

then

$$\int_0^\infty x^{-r}(f_1(x))^p\,dx \leqq K\int_0^\infty x^{-r}(xf(x))^p\,dx, \tag{8.1.1}$$

$$\int_0^\infty x^{r-2}(f_2(x))^p\,dx \leqq K\int_0^\infty x^{r-2}(f(x))^p\,dx, \tag{8.1.2}$$

with $K = K(p, r)$, whenever the integrals on the right are finite.

These are known theorems.[26] The cases we require are $r = p$ and $r = 2$.

[25] Hardy, Littlewood, and Pólya (**10**), Theorem 319.

[26] For (8.1.1) see Hardy, Littlewood, and Pólya (**10**), Theorem 330. The second inequality is not stated explicitly in the book, but will be found in Hardy (**1**).

LEMMA β. *If $1 < p \leqq 2$, and $f(x)$ is $L^p(0, \infty)$, then*

$$F(x) = \int_0^\infty f(t) \begin{matrix}\cos\\ \sin\end{matrix}\, xt\, dt = \operatorname*{l.i.m.}_{a\to\infty} \int_0^a f(t) \begin{matrix}\cos\\ \sin\end{matrix}\, xt\, dt$$

exists, for almost all x, as a limit in mean with index p', and

$$\int_0^\infty x^{p-2} \mid F(x) \mid^p dx \leqq K(p) \int_0^\infty \mid f(x) \mid^p dx . \tag{8.1.3}$$

LEMMA γ. *If $p \geqq 2$, and $x^{(p-2)/p} f(x)$ belongs to $L^p(0, \infty)$, then*

$$F(x) = \int_0^\infty f(t) \begin{matrix}\cos\\ \sin\end{matrix}\, xt\, dt = \operatorname*{l.i.m.}_{a\to\infty} \int_0^a f(t) \begin{matrix}\cos\\ \sin\end{matrix}\, xt\, dt$$

exists, for almost all x, as a limit in mean with index p, and

$$\int_0^\infty \mid F(x) \mid^p dx \leqq K(p) \int_0^\infty x^{p-2} \mid f(x) \mid^p dx . \tag{8.1.4}$$

For these two theorems see Hardy and Littlewood (**3**), Theorems 13 and 14.

8.2. LEMMA δ. *Let*

$$\psi(x) = x^{1/p} \int_0^1 (1 - u)^{-1/p'} \cos xu\, du . \tag{8.2.1}$$

Then the result of Lemma β remains true when $\psi(xt)$ is substituted for cos *xt or* sin *xt*.

It is easily verified by standard methods[27] that $\psi(x)$ is regular for $0 < x < \infty$, that

$$\psi(x) \backsim px^{1/p}$$

for small positive x, and that

$$\psi(x) = \Gamma\left(\frac{1}{p}\right) \cos\left(x - \frac{\pi}{2p}\right) + O\left(\frac{1}{x}\right)$$

for large positive x. Hence

$$\psi(x) = C \cos\left(x - \frac{\pi}{2p}\right) + R(x) , \tag{8.2.2}$$

where

$$\mid R(x) \mid < K\ (0 < x \leqq 1) , \qquad \mid R(x) \mid < \frac{K}{x}\ (x > 1) , \tag{8.2.3}$$

and $C = C(p)$, $K = K(p)$.

[27] The simplest method for finding an asymptotic expansion for $\psi(x)$ is to apply Cauchy's Theorem to

$$\int (1 - u)^{-1/p'} e^{ixu}\, du$$

and the rectangle $(0, 1, 1 + i\infty, i\infty)$.

It is enough to prove the result on the hypothesis that $f(x) = 0$ for $x > c$, when

$$F(x) = \int_0^\infty f(t)\, \psi(xt)\, dt$$

exists, for all x, as a Lebesgue integral; the proof may then be completed as at the end of §7.2.

Now in this case[28]

$$F(x) = C \int_0^\infty f(t) \cos\left(xt - \frac{\pi}{2p}\right) dt + \int_0^{1/x} f(t)\, R(xt)\, dt + \int_{1/x}^\infty f(t)\, R(xt)\, dt$$
$$= F_1(x) + F_2(x) + F_3(x)\,,$$

say; and it is sufficient to show that F_1, F_2, and F_3 satisfy inequalities of the type (8.1.3). This is true of F_1, by Lemma β. Next

$$| F_2 | \leqq K \int_0^{1/x} | f(t) |\, dt = K f_1\!\left(\frac{1}{x}\right),$$

say; and so

$$\int_0^\infty x^{p-2} | F_2 |^p\, dx \leqq K \int_0^\infty x^{p-2} \left(f_1\!\left(\frac{1}{x}\right)\right)^p dx$$
$$= K \int_0^\infty x^{-p} (f_1(x))^p\, dx \leqq K \int_0^\infty | f(x) |^p\, dx\,,$$

by (8.1.1), with $r = p$. Finally

$$| F_3 | \leqq \frac{K}{x} \int_{1/x}^\infty \frac{| f(t) |}{t}\, dt = \frac{K}{x} f_2\!\left(\frac{1}{x}\right),$$

say, and

$$\int_0^\infty x^{p-2} | F_3 |^p\, dx \leqq K \int_0^\infty x^{-2} \left(f_2\!\left(\frac{1}{x}\right)\right)^p dx$$
$$= K \int_0^\infty (f_2(x))^p\, dx \leqq K \int_0^\infty | f(x) |^p\, dx\,,$$

by (8.1.2), with $r = 2$.

8.3. Lemma ϵ. *Let*

$$\chi(x) = x^{1/p'} \int_0^1 (1 - u)^{-1/p} \cos xu\, du\,.$$

Then the result of Lemma γ remains true when $\chi(xt)$ is substituted for cos xt *or* sin xt.

[28] Our argument is suggested by one used by Titchmarsh (**14**) for a different purpose.

Here

$$\chi(x) = C \cos\left(x - \frac{\pi}{2p'}\right) + R(x) ,$$

where $R(x)$ again satisfies (8.2.3). Arguing as before, we obtain

$$F(x) = F_1(x) + F_2(x) + F_3(x) ,$$

where

$$\int_0^\infty | F_1(x) |^p \, dx \leqq K(p) \int_0^\infty x^{p-2} | f(x) |^p \, dx$$

and

$$| F_2(x) | \leqq K f_1\left(\frac{1}{x}\right), \qquad | F_3(x) | \leqq \frac{K}{x} f_2\left(\frac{1}{x}\right).$$

We have now

$$\int_0^\infty | F_2(x) |^p \, dx \leqq K \int_0^\infty \left(f_1\left(\frac{1}{x}\right)\right)^p dx$$

$$= K \int_0^\infty x^{-2} (f_1(x))^p \, dx \leqq K \int_0^\infty x^{p-2} | f(x) |^p \, dx$$

by (8.1.1), with $r = 2$; and

$$\int_0^\infty | F_3(x) |^p \, dx \leqq K \int_0^\infty x^{-p} \left(f_2\left(\frac{1}{x}\right)\right)^p dx$$

$$= K \int_0^\infty x^{p-2} (f_2(x))^p \, dx \leqq K \int_0^\infty x^{p-2} | f(x) |^p \, dx ,$$

by (8.1.2) with $r = p$. The result follows as before.

9. Proof of Theorem 13

9.1. (1) Suppose that $1 < p \leqq 2$, that $x^{-1/p} f(x)$ is L^p, and, in the first instance, that $f(x) = 0$ for $x > c$.

Let

$$w(x) = \int_0^\infty t^{-1/p} f(t) \psi(xt) \, dt ,$$

where ψ is defined as in §8.2. Then

$$\int_0^x (x - y)^{-1/p} w(y) \, dy = \int_0^x (x - y)^{-1/p} \, dy \int_0^\infty t^{-1/p} f(t) \psi(yt) \, dt$$

$$= \int_0^\infty t^{-1/p} f(t) \, dt \int_0^x (x - y)^{-1/p} \psi(yt) \, dy$$

(by absolute convergence). The inner integral is

$$\int_0^x (x-y)^{-1/p}(yt)^{1/p}\,dy \int_0^1 (1-u)^{-1/p'} \cos ytu\,du$$

$$= t^{1/p}\int_0^x (x-y)^{-1/p}\,dy \int_0^y (y-v)^{-1/p'} \cos tv\,dv$$

$$= t^{1/p}\int_0^x \cos tv\,dv \int_v^x (x-y)^{-1/p}(y-v)^{-1/p'}\,dy = \pi \operatorname{cosec}\frac{\pi}{p}\cdot t^{1/p}\frac{\sin xt}{t};$$

so that

$$s(x) = \int_0^\infty f(t)\frac{\sin xt}{t}\,dt = K(p)\int_0^x (x-y)^{-1/p} w(y)\,dy$$

is, apart from a factor K, the $(1/p')$-th integral of $w(x)$.

It follows from Theorem 7 and Lemma δ that

$$\int_0^\infty \frac{|s(x)|^p}{x}\,dx \leqq K\int_0^\infty x^{p-2}|w(x)|^p\,dx \leqq K\int_0^\infty \frac{|f(x)|^p}{x}\,dx.$$

This is the result of Theorem 13, when $f(x)$ is 0 for large x; and the full result follows as in §7.2.

(2) If $p \geqq 2$ we write

$$w(x) = \int_0^\infty t^{-1/p'} f(t)\chi(xt)\,dt,$$

where χ is defined as in §8.3. We again suppose, in the first instance, that $f(x) = 0$ for large x. Then $s(x)$ is substantially the $(1/p)$-th integral of $w(x)$, and

$$\int_0^\infty \frac{|s(x)|^p}{x}\,dx \leqq K\int_0^\infty |w(x)|^p\,dx \leqq K\int_0^\infty x^{p-2}\left(\frac{|f(x)|}{x^{1/p'}}\right)^p dx$$

$$= K\int_0^\infty \frac{|f(x)|^p}{x}\,dx,$$

by Theorem 6 and Lemma ϵ. The proof is then completed as before.

10. Concluding remarks

10.1. We conclude with a few miscellaneous comments.

(1) The result of Theorem 10 becomes false for $p = 1$.

It is plain that $\sum n^{-1}|s_n| < \infty$ implies $\sum n^{-1}|a_n| < \infty$ and so

$$\sum \frac{|s_n - \frac{1}{2}a_n|}{n} < \infty.$$

Also

$$s_n - \frac{1}{2}a_n = \frac{2}{\pi}\int_0^{\frac{1}{2}\pi} \frac{1}{2}\cot\frac{1}{2}t\, f(t)\sin nt\,dt$$

is the Fourier sine coefficient of the function $\frac{1}{2}\cot\frac{1}{2}t\,f(t)$. Hence, if the result were true for $p = 1$, it would also be true that, if $g(t)$ is odd and integrable, and

$$g(t) \sim \sum b_n \sin nt,$$

then

$$\sum \frac{|b_n|}{n} < \infty.$$

But

$$g(t) = 2\sum \frac{\sin nu}{\log n} \sin nt = h(t-u) - h(t+u),$$

where

$$h(t) = \sum \frac{\cos nt}{\log n},$$

is integrable, for any u; while

$$\sum \frac{|\sin nu|}{n \log n}$$

is generally divergent.

Thus

$$\int_0 \frac{|f(t)|}{t}\,dt < \infty$$

implies $s_n \to 0$ (by Dini's convergence criterion), but not the convergence of $\sum n^{-1}|s_n|$. When $p > 1$ the situation is reversed:

$$\int_0 \frac{|f(t)|^p}{t}\,dt < \infty \tag{10.1.1}$$

implies the convergence of $\sum n^{-1}|s_n|^p$, but does not imply $s_n \to 0$. For (10.1.1) is satisfied whenever

$$f(t) = O\left(\left(\log \frac{1}{t}\right)^{-1}\right),$$

and this is not a sufficient condition for convergence of the Fourier series.[29]

10.2. (2) It is instructive to contrast our results with the much simpler results for the Cesàro mean σ_n of the Fourier series.

If (10.1.1) is satisfied then, à fortiori,

$$\int_0^t |f(u)|^p\,du = o(t)$$

[29] See Hardy and Littlewood (**6**), 47; Zygmund (**15**), 31, 174.

and this is, for $p \geqq 1$, a sufficient condition that $\sigma_n \to 0$. Also (10.1.1) implies

$$\sum \frac{|\sigma_n|^p}{n} < \infty. \tag{10.2.1}$$

When $p > 1$, this is a corollary of Theorem 10; but it is true for $p \geqq 1$, and may be proved much more simply. For

$$|\sigma_n| \leqq A \int_0^\pi |f(t)| \frac{\sin^2 nt}{nt^2}\, dt \leqq An \int_0^{1/n} |f(t)|\, dt + \frac{A}{n} \int_{1/n}^\pi \frac{|f(t)|}{t^2}\, dt, \tag{10.2.2}$$

where the A are constants; and (10.2.1) is an easy deduction.

Consider, for example, the first term

$$An \int_0^{1/n} |f(t)|\, dt = An f_1\left(\frac{1}{n}\right)$$

on the right of (10.2.2). The contribution of this to (10.2.1) does not exceed

$$K \sum n^{p-1} \left(f_1\left(\frac{1}{n}\right)\right)^p \leqq K \sum \int_n^{n+1} x^{p-1} \left(f_1\left(\frac{1}{x}\right)\right)^p dx \leqq K \int_0^\infty x^{p-1} \left(f_1\left(\frac{1}{x}\right)\right)^p dx$$
$$= K \int_0^\infty x^{-p-1} f_1^p(x)\, dx.$$

We now require the inequality

$$\int_0^\infty x^{-p-1} (f_1(x))^p\, dx \leqq K \int_0^\infty x^{-1} (f(x))^p\, dx,$$

which is a case of (8.1.1) when $p > 1$ and may be verified independently when $p = 1$.

The second term in (10.2.2) may be disposed of similarly.

10.3. (3) Theorem 10 has a 'transform,' viz.

THEOREM 14. *If $p > 1$ and*

$$\sum n^{p-1} |b_n|^p < \infty,$$

then there is an odd function $g(x)$ whose Fourier series is

$$\sum b_n \sin nx,$$

and

$$\left(\int_0^\pi \frac{|g(x)|^k}{x}\, dx\right)^{1/k} \leqq K(p, k) \left(\sum n^{p-1} |b_n|^p\right)^{1/p}$$

for $k \geqq p$.

This may be proved independently, or (when $k = p$) deduced from Theorem 13; and there is a simple proof similar to that of §7.2 when $k = p = 2$.

The corresponding theorem for cosine series is false. If $b_1 = 0$, and $b_n = (n \log n)^{-1}$ for $n > 1$, $k = p = 2$, then

$$\sum nb_n^2 = \sum \frac{1}{n(\log n)^2} < \infty;$$

but

$$f(t) = \sum \frac{\cos nt}{n \log n} \sim \int_t^\pi \sum \frac{\sin nu}{\log n}\, du \sim \int_t^\pi \frac{du}{u \log (1/u)} \sim \log \log \frac{1}{t}$$

for small positive t.

(4) Theorem 13 is equivalent to

THEOREM 15. *The bilinear integral form*

$$\int_0^\infty \int_0^\infty \frac{\sin xy}{x^{1/p'} y^{1/p}} a(x) b(y)\, dx\, dy$$

is bounded in space $[p, p']$: *i.e.,*

$$\left| \int_0^X \int_0^Y \frac{\sin xy}{x^{1/p'} y^{1/p}} a(x) b(y)\, dx\, dy \right| \leqq K(p) \left(\int_0^\infty |a(x)|^p\, dx \right)^{1/p} \left(\int_0^\infty |b(y)|^{p'}\, dy \right)^{1/p'}$$

for all $a(x)$, $b(y)$, X, Y.

The form

$$\sum\sum \frac{\sin mn}{m^{1/p'} n^{1/p}} a_m b_n$$

is not bounded in $[p, p']$, since (*e. g.*)

$$\sum_m \frac{|\sin mn|^{p'}}{m} = \infty$$

for every n.[30]

REFERENCES

1. G. H. Hardy, *Notes on some points in the integral calculus* (LXIV), Messenger of Math., vol. 57 (1928), pp. 12–16.
2. G. H. Hardy and J. E. Littlewood, *Sur la série de Fourier d'une fonction à carré sommable*, Comptes Rendus, vol. 156 (1913), pp. 1307–1309.
3. ——— *Some new properties of Fourier constants*, Math. Ann., vol. 97 (1926), pp. 159–209.
4. ——— *Notes on the theory of series* (IV): *On the strong summability of Fourier series*, Proc. London Math. Soc., (2), vol. 26 (1927), pp. 273–286.
5. ——— *Notes on the theory of series* (V): *On Parseval's theorem*, Proc. London Math. Soc., (2), vol. 26 (1927), pp. 287–294.
6. ——— *Some new convergence criteria for Fourier series*, Annali Scuola Norm. Sup. Pisa, vol. 3 (1934), pp. 43–62.
7. ——— *An inequality*. Math. Zeitschr., vol. 40, (1935), pp. 1–40.
8. ——— *The strong summability of Fourier series*, Fundamenta Math., vol. 25 (1935), pp. 162–189.
9. G. H. Hardy, J. E. Littlewood, and G. Pólya, *The maximum of a certain bilinear form*, Proc. London Math. Soc., (2), vol. 25 (1926), pp. 265–282.

[30] See Hardy, Littlewood and Pólya (**10**), Theorem 289.

10. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1934.
11. F. Hausdorff, *Eine Ausdehnung des Parsevalschen Satzes über Fourierreihen*, Math. Zeitschr., vol. 16 (1923), pp. 163–169.
12. E. Lindelöf, *Le calcul des Résidus*, Paris, 1905.
13. M. Riesz, *Sur les fonctions conjuguées*, Math. Zeitschr., vol. 27 (1928), pp. 218–244.
14. E. C. Titchmarsh, *A note on Hankel transforms*, Journal London Math. Soc., vol. 1 (1926), pp. 195–196.
15. A. Zygmund, *Trigonometrical Series*, Warszawa-Lwów, 1935.

CAMBRIDGE UNIVERSITY.

SUR QUELQUES DÉFINITIONS POSSIBLES DE L'INTÉGRALE DE STIELTJES

PAR MAURICE FRÉCHET

Introduction. La définition la plus utile de l'intégrale de Stieltjes $\int_V F(x)dv(x)$ (où V est le domaine d'intégration) est celle qui concerne le cas où $F(x)$ est continu, $v(x)$ à variation bornée et où V est un segment fini. MM. F. Riesz et Lebesgue ont étendu cette définition, en la modifiant, au cas où $f(x)$ n'est pas continue.

En Calcul des Probabilités, deux extensions plus modestes paraissent seulement désirables. Celle concernant $F(x)$, où l'on suppose $F(x)$ monotone et qui trouve son application dans la détermination de la fonction des probabilités totales de la somme de deux variables indépendantes, forme l'objet de la Première Partie de ce mémoire. Les résultats obtenus prolongent en les complétant—en particulier, en adjoignant à leurs conditions suffisantes des conditions nécessaires—certains résultats antérieurs de MM. Lebesgue, Steffensen et Glivenko.

La seconde extension, concernant le cas où $F(x)$ est supposé continu mais où V est illimité, est le sujet traité dans la Seconde Partie de ce mémoire. Elle se trouve utile dans la détermination de la valeur et des propriétés de la valeur moyenne d'une fonction continue $f(X)$ d'une variable aléatoire X.

Les définitions de $\int_V f(x)\, dC(x)$.

Première partie. Cas où $f(x)$ est monotone

Extension de la définition de l'intégrale de Stieltjes. Quand $f(x)$ et $C(x)$ sont deux fonctions définies sur un intervalle fini (α, β), la première continue, l'autre monotone, on démontre que la somme

$$\sigma = \sum_i f(\xi_i)\, [C(x_i) - C(x_{i-1})]$$

tend vers une limite déterminée I quand, en prenant arbitrairement $\alpha = x_0 \leqq \xi_1 \leqq x_1 \leqq \xi_2 \leqq \cdots \leqq x_{n-1} \leqq \xi_n \leqq x_n = \beta$, la plus grande δ des différences $x_i - x_{i-1}$ tend vers zéro. Et on pose $I = \int_\alpha^\beta f(x)dC(x)$. On a étendu cette définition à des cas beaucoup plus généraux. Il nous sera seulment utile ici de considérer le cas où f, continue ou non, est monotone comme $C(x)$. Nous verrons quelle restriction il y a à faire pour que l'extension soit possible.

Received March 2, 1936.

Décomposition de l'intégrale. Pour simplifier l'étude, on peut utiliser la propriété suivante: si $f(x)$ est monotone, $f(x)$ est la somme de sa "fonction des sauts" $S(x)$ et d'une fonction continue $\varphi(x)$. On a alors $\sigma = \sigma_1 + \tau$, où σ_1, τ sont formées comme σ, mais à partir de φ et S au lieu de f. Dans ce qui suit, nous pourrons envisager le cas général d'un segment V d'intégration, fini ou non, soit (α, β), $(\alpha, +\infty)$, $(-\infty, \beta)$, $(-\infty, +\infty)$ pourvu que nous supposions *les fonctions f et C bornées*—ce qui a lieu nécessairement (si on les suppose finies en chaque point) quand le segment V est fini. Quand V est infini, la suite des x_i sera illimitée dans un sens ou dans les deux. En appelant l, L les bornes de C; h, H celles de f, la fonction $f_1(x) = f(x) - h$ a évidemment les mêmes sauts. On posera alors dans tous les cas

$$S(x) = \sum_{r_i \leqq x} s_i' + \sum_{r_i < x} s_i'' \tag{1}$$

avec

$$s_i' = f(r_i) - f(r_i - 0)\,, \tag{2}$$

$$s_i'' = f(r_i + 0) - f(r_i)\,, \tag{3}$$

$r_1, r_2, \cdots$ étant les points de discontinuité de f ou encore de f_1. Il est clair que $0 \leqq S(x) \leqq H - h$ et que $S(x)$ est non-décroissant comme $f(x)$. Si f est borné, φ l'est donc aussi. Quand φ est continu *et borné* et C monotone *et borné*, la somme $\sigma_1 = \Sigma\, \varphi(\xi_i)[C(x_i) - C(x_{i-1})]$ est une somme d'un nombre fini de termes ou une série absolument convergente. Et la méthode classique montre que σ_1 a une limite unique quand $\delta \to 0$:

$$\lim \sigma_1 = J = \int_V \varphi(x)\, dC(x)\,.$$

Pour établir l'existence de la limite de σ, il suffit donc d'établir l'existence de

$$K = \lim \tau, \text{ avec } \tau = \sum_i S(\xi_i)\, [C(x_i) - C(x_{i-1})]\,.$$

S'il en est ainsi, on pourra encore poser $\int_a^b f(x)\, dC(x) = \lim \sigma = J + K$. Or, en supposant, par exemple, $f(x)$ non-décroissante aussi que $C(x)$, on aura $t \leqq \tau \leqq T$ avec

$$\tau = \sum_i S(x_{i-1})\, [C(x_i) - C(x_{i-1})]\,; \qquad T = \sum_i S(x_i)\, [C(x_i) - C(x_{i-1})]\,. \tag{4}$$

Soient t_1 et T_2 les valeurs de t et T pour deux modes de divisions quelconques, t' et T' leurs valeurs pour le mode de division obtenu en combinant les précédents. On a évidemment, si, par exemple, $f(x)$ et $C(x)$ sont non-décroissants,

$$t_1 \leqq t' \leqq T' \leqq T_2\,.$$

Ainsi on a d'abord $t_1 \leqq T_2$ et les t_1 ont une borne supérieure finie m, les T_2 une borne inférieure M avec $m \leqq M$.

Existence des limites de t et de T. Cherchons d'abord ce que deviennent t et T quand $\delta \to 0$. La série $S(x)$ est convergente, et on a

$$T = \sum_i \Big[\sum_{r_j \leqq x_i} s'_j + \sum_{r_j < x_i} s''_j \Big] [C(x_i) - C(x_{i-1})] .$$

Cette série double à termes $\geqq 0$ est convergente, et on peut écrire

$$T = \sum_j s'_j \Big\{ \sum_{r_j \leqq x_i} [C(x_i) - C(x_{i-1})] \Big\} + \sum_j s''_j \Big\{ \sum_{r_j < x_i} [C(x_i) - C(x_{i-1})] \Big\} ,$$

$$T = \sum_j s'_j [L - C(x'_j)] + \sum_j s''_j [L - C(x''_j)] , \tag{5}$$

x'_j et x''_j étant des points pris dans la suite des x_i, soient $x_{\alpha-1}$ et $x_{\beta-1}$, tels que

$$x_{\alpha-1} < r_j \leqq x_\alpha ; \qquad x_{\beta-1} \leqq r_j < x_\beta . \tag{6}$$

Que se passe-t-il pour T quand $\delta \to 0$? Les deux points $x_{\alpha-1}$, $x_{\beta-1}$ tendent vers r_j; $x_{\alpha-1}$ tend vers r_j par valeurs inférieures, donc le terme $C(x'_j)$ de T tend vers $C(r_j - 0)$. Quant à $C(x''_j)$, sa limite dépend de la manière dont $x_{\beta-1}$ tend vers r_j. Si, à partir d'une valeur assez petite de δ, la division D formée des points x_i ne comprend pas r_j, $C(x''_j) = C(x_{\beta-1}) \to C(r_j - 0)$. Si r_j appartient à D pour ρ assez petit $C(x''_j) \to C(r_j)$. Ainsi la plus petite des limites et la plus grande des limites de T quand $\delta \to 0$ sont[1]

$$\begin{aligned} A &= \sum_j s'_j [L - C(r_j - 0)] + \sum_j s''_j [L - C(r_j)] , \\ A' &= \sum_j s'_j [L - C(r_j - 0)] + \sum_j s''_j [L - C(r_j - 0)] . \end{aligned} \tag{7}$$

On a

$$A' - A = \sum_j s''_j [C(r_j) - C(r_j - 0)] . \tag{8}$$

On verrait de même que la plus grande et la plus petite des limites de t quand δ tend vers zéro sont

$$\begin{aligned} a &= \sum_j s'_j [L - C(r_j)] + \sum_j s''_j [L - C(r_j + 0)] , \\ a' &= \sum_j s'_j [L - C(r_j + 0)] + \sum_j s''_j [L - C(r_j + 0)] , \end{aligned} \tag{9}$$

et on a

$$a - a' = \sum_j s'_j [C(r_j + 0) - C(r_j)] . \tag{10}$$

[1] A vrai dire, ceci montre seulement que les termes de T tendent respectivement suivant le cas vers les termes correspondants de A ou de A'. Mais les termes de T, A, A' sont inférieurs à ceux de la série convergente et indépendante de δ: $\Sigma\, s'_j[C(b) - C(a)] + \Sigma\, s''_j [C(b) - C(a)]$, ce qui permet de compléter la démonstration.

Si donc f et C sont deux fonctions monotones bornées quelconques, τ et par suite σ n'ont pas nécessairement des limites déterminées, *l'intégrale de Stieltjes* $\int_V f(x)dC(x)$ *peut ne pas exister.* Pour assurer l'unicité de la limite de T et de la limite de t deux moyens se présentent. Ou bien *restreindre l'arbitraire de la fonction f* ou bien *restreindre celui des divisions D.*

Le premier moyen a été choisi par M. Steffensen[2] dans une étude où, d'ailleurs, en vue des applications actuarielles, il lui a paru suffisant de supposer que f n'a qu'un nombre fini de discontinuités et de prendre $\xi_i = (x_i + x_{i-1})/2$. Nous allons montrer que la condition suffisante qu'il obtient sous ces restrictions pour l'unicité de σ l'est en même temps pour celles de t et T dans notre cas plus général, et en outre qu'elle est nécessaire pour l'unicité de t et celle de T.

Pour que T et t aient, tous deux, des limites uniques, *il faut* et *il suffit,* d'après ce qui précède, que $A' - A = a - a' = 0$. Les expressions ci-dessus de $A' - A$ et $a - a'$ étant formées de termes $\geqq 0$, ces termes devront être tous nuls. En se reportant aux expressions (8), (10), on voit que cette condition peut s'exprimer ainsi: *en tout point de discontinuité commun à f et C, f(x) et C(x) doivent être continues à la fois d'un certain côté de ce point,* ce côté pouvant d'ailleurs varier avec le point considéré. C'est la condition imposée à priori par M. Steffensen et dont nous trouvons qu'elle est nécessaire et suffisante pour que T et t n'ait chacun qu'une limite, indépendante du choix de la suite de divisions D pourvu que δ tende vers zéro.

Mais on peut, avec M. Lebesgue, supprimer cette restriction sur f en introduisant une restriction sur les divisions D. Le raisonnement précédent montre en effet que si la condition de M. Steffensen n'est pas réalisée, pour qu'une suite de divisions D fournisse une limite unique de T et une limite unique de t, il faut et il suffit que, si en un point de discontinuité x commun à f et C, il n'y a aucun côté de x où $f(x)$ et $C(x)$ soient continues toutes les deux, ce point x doit, à partir d'une valeur assez petite de δ, appartenir toujours ou n'appartenir jamais à la suite des D.

A cet effet, il suffit, par exemple, que tout point de discontinuité de $C(x)$ appartienne à la suite des divisions D pour δ assez petit. C'est la condition suffisante indiquée par M. Lebesgue[3] dans un cas beaucoup plus général, celui où $C(x)$ étant a variation bornée, $f(x)$ est seulement supposée bornée.

D'après ce qui précède, il y a ici d'autres choix possibles et moins restreints de la suite des D. Par exemple, il suffit que tout point de discontinuité *commun* à $f(x)$ et à $C(x)$ appartienne à la suite des divisions à partir d'une valeur correspondante assez petite de δ.

Il y a des choix encore moins restreints de la suite des D, qui donneraient chacun à T une limite unique mais distincte de celle qui correspondrait au choix précédent. Pour éviter cette indétermination nous allons faire intervenir l'unicité de la limite, non seulement de T et de t, mais encore de τ.

[2] *On Stieltjes' integral and its application to actuarial questions,* Journal of the Institute of Actuaries, vol. 63 (1932), p. 447.

[3] *Leçons sur l'Intégration,* 2ième édition, 1928, p. 272.

T et t sont deux valeurs possibles de τ et l'on a $t \leqq \tau \leqq T$. Donc pour que τ ait pour une suite de divisions D une limite unique indépendante du choix des ξ_i dans les intervalles (x_{i-1}, x_i), il faut et il suffit: (i) que T et t aient chacun une limite déterminée; (ii) que les limites de T et de t soient égales. La différence des limites de T et t est en tout cas au moins égale à

$$A - a = \sum_j s'_j [C(r_j) - C(r_j - 0)] + \sum_j s''_j [C(r_j + 0) - C(r_j)]. \tag{11}$$

Dès lors, pour que τ ait une limite indépendante du choix des ξ_i, *il faut* que chacun des termes (qui sont tous $\geqq 0$) de $A - a$ soit nul; c'est à dire:

Condition (N). *Il faut que, de chaque côté de chaque point, l'une au moins des fonctions f et C soit continue, sans qu'il s'agisse nécessairement de la même fonction quand le côté ou le point change.*

La condition (N) est nécessaire pour que σ converge vers une limite indépendante du choix des ξ_i quand on considère une suite déterminée de divisions D telle que δ tende vers zéro. Mais il faut aussi que T et t convergent chacun vers une limite unique. Et comme ces limites sont respectivement comprises entre A et A', a et a', et que la condition (N) assure seulement l'égalité $A = a$ avec $A' \geqq A = a \geqq a'$, il faut que T et t tendent précisément vers A et a. C'est à dire, comme il ressort de l'analyse faite plus haut, que tout point r_j doit appartenir à la suite des D à partir d'une valeur assez petite de δ, si en ce point

$$s''_j [C(r_j) - C(r_j - 0)] + s'_j [C(r_j + 0) - C(r_j)] \neq 0.$$

Or si la condition (N) est réalisée, on a

$$s'_j [C(r_j) - C(r_j - 0)] + s''_j [C(r_j + 0) - C(r_j)] = 0,$$

et par suite, en ajoutant,

$$(s'_j + s''_j)[C(r_j + 0) - C(r_j - 0)] \neq 0,$$

ou, puisque $s'_j + s''_j \neq 0$,

$$C(r_j + 0) - C(r_j - 0) \neq 0.$$

Les points r_j où cette inégalité est vérifiée sont les points de discontinuité communs à f et C. Le choix de divisions que nous avions donc indiqué plus haut comme simplement suffisant pour assurer l'unicité de chacune des limites de T et t, quand on ne sait rien sur la réalisation de (N), devient donc nécessaire quand (N) étant supposée réalisée, on veut assurer l'égalité des limites de T et de t et par suite l'indépendance pour la limite de τ—et par suite de σ—du choix des ξ_i dans les (x_{i-1}, x_i).

En résumé: pour que σ tende vers une limite indépendante du choix des ξ_i dans les (x_{i-1}, x_i), quand on considère les valeurs de σ correspondant à une suite de divisions D convenablement choisie et telle que δ tende vers zéro, il faut et il suffit que (i) la condition (N) soit realisée, (ii) tout point de discontinuité commun à $f(x)$ et $C(x)$ appartienne à la suite des D à partir d'une valeur correspondante suffisamment petite de δ.

Autre définition de l'intégrale. On retrouve encore la condition nécessaire (N) mais qui devient aussi suffisante à elle seule quand on se place au point de vue de M. Glivenko.[4] Celui-ci étend au cas de l'intégrale de Stieltjes la définition des intégrales de Darboux et il généralise la définition de l'intégrale ordinaire en exigeant que leurs valeurs soient égales. Plus précisément, d'après M. Glivenko, l'intégrale de Stieltjes $\int_V S(x)dC(x)$ existe quand les bornes M et m des sommes de Darboux sont égales, et leur valeur commune est la valeur de l'intégrale. Mais les intervalles de variation de x et les variations correspondantes de $C(x)$ sont supposés définis de façon convenable.[4]

Sans avoir besoin de faire varier δ, on voit en comparant les expressions (5) et (7) de T et de A, qu'on a toujours $T \geqq A$. D'où $M \geqq A$. Or on vient de voir que pour certaines suites de divisions, T tend vers A; comme sa limite ne peut être que $\geqq M$, on a donc $A \geqq M$. Finalement $M = A$ et de même $m = a$. Dès lors, pour que $M = m$, il faut et il suffit que $A - a = 0$ et on retombe bien sur la condition ci-dessus.

D'ailleurs, écrivons

$$\sum_i f(x_i)[C(x_i) - C(x_{i-1})] = \sum_i \varphi(x_i)[C(x_i) - C(x_{i-1})] + \sum_i S(x_i)[C(x_i) - C(x_{i-1})].$$

Si M' et M'' sont les bornes inférieures des deux premières sommes, on voit qu'on a

$$\sum_i f(x_i)[C(x_i) - C(x_{i-1})] \geqq M'' + M,$$

d'où $M' \geqq M'' + M$. D'autre part, pour tout $\epsilon > 0$, il existe une division D_ϵ pour laquelle $T < M + \epsilon$, et, puisque φ est continue et bornée, un nombre ρ tel que pour $\delta < \rho$, on ait

$$\sum_i \varphi(x_i)[C(x_i) - C(x_{i-1})] < M'' + \epsilon.$$

Adjoignons aux points de D_ϵ des points choisis de sorte que pour la division D obtenue, δ devienne $< \rho$. Cette opération ne peut que diminuer la seconde somme. Pour cette division D, la deuxième et la troisième sommes sont donc respectivement inférieures à $M'' + \epsilon$, $M + \epsilon$. Dès lors, la première qui est $\geqq M'$ sera $\leqq M'' + M + 2\epsilon$. Des inégalités ainsi obtenues

$$M'' + M \leqq M' \leqq M'' + M + 2\epsilon,$$

vérifiées quel que soit ϵ, on tire $M' = M + M''$. Avec des notations analogues, on établirait de même que $m' = m + m''$, d'où $M' - m' = M - m + M'' - m'' = M - m$, puisque, φ étant continu et borné, on a, comme on sait, $M'' = m''$.

[4] *Sur les sommes de variables aléatoires.* Ce travail (en français) qui doit être imprimé dans les travaux du Séminaire des Probabilités de l'Université de Moscou et dont M. Glivenko a bien voulu me communiquer une copie, sera reproduit dans l'ouvrage (en russe) intitulé *Intégrale de Stieltjes*, par Valère Glivenko, 1936.

Dès lors, pour que $M' = m'$, il faut et il suffit que $M = m$ ou encore $A = a = 0$, ce qui conduit encore à la condition déjà signalée plus haut. Celle-ci est donc la condition nécessaire et suffisante pour que $\int_V f(x)dC(x)$ existe au sens de M. Glivenko dans le cas actuel.

Revenons à la limite de σ. On sait que σ_1 a une limite unique indépendante du choix des ξ_i dans les segments (x_{i-1}, x_i) et de la suite des divisions D, pourvu que δ tende vers zéro. Pour qu'il en soit de même de $\sigma = \sigma_1 + \tau$, il faut donc et il suffit qu'il en soit ainsi pour τ. La condition cherchée résultera de la combinaison de la dernière condition obtenue[5] et de la condition de M. Steffensen. C'est à dire que $f(x)$ et $C(x)$ ne devront avoir aucun point de discontinuité en commun.

Remarque. Dans tout ce qui précède nous avons supposé, par exemple, f et C tous deux non-décroissants. On ramène à ce cas celui où f ou C ou tous les deux seraient non-croissants en remplaçant f ou C par $-f$ ou $-C$.

En résumé: Soient $f(x)$, $C(x)$ deux fonctions monotones sur le segment ab, *pour que l'intégrale* $\int_a^b f(x)dC(x)$ *existe, il faut et il suffit*

I. Si on la définit comme la valeur commune de la borne supérieure de

$$\sum_i f(x_{i-1})[C(x_i) - C(x_{i-1})]$$

et de la borne inférieure de

$$\sum_i f(x_i)[C(x_i) - C(x_{i-1})],$$

que, de chaque côté de chaque point x, l'une ou l'autre des deux fonctions $f(x)$ et $C(x)$ soit continue;

II. Si on la définit comme limite unique de

$$\sigma = \sum_i f(\xi_i)[C(x_i) - C(x_{i-1})]$$

pour un choix quelconque des ξ_i, mais pour une suite de divisions *convenables* telle que δ tende vers zéro, *que, de chaque côté de chaque point x, l'une ou l'autre des deux fonctions $f(x)$ et $C(x)$ soit continue et alors que la suite des divisions considérées comprenne chaque point de discontinuité commun à $C(x)$ et à $f(x)$ à partir d'une valeur de δ assez petite* (pouvant éventuellement varier avec ce point);

III. Si on la définit comme limite unique de σ quels que soient le choix des ξ_i dans les segments (x_{i-1}, x_i) et aussi quelle que soit la suite des divisions D pourvu que δ tende vers zéro, que *$f(x)$ et $C(x)$ n'aient aucun point de discontinuité en commun.*

[5] Car celle ci, nécessaire en général, devient suffisante quand on la combine avec celle de M. Steffensen, puisque dans ce cas la différence des limites de T et t non seulement est $\geqq A - a$, mais, étant $\leqq A' - a'$, est alors égale à $A - a = A' - a'$.

Calcul de $\int_V S(x)dC(x)$. Chemin faisant nous avons obtenu tout ce qu'il faut pour calculer l'expression de $\int_V S(x)dC(x)$. Quand elle existe au moins selon les definitions I ou II, elle est égale à $A = a$, donc aussi à

$$\frac{A+a}{2} = \sum_j s_j'\left[L - \frac{C(r_j) + C(r_j - 0)}{2}\right] + \sum_j s_j''\left[L - \frac{C(r_j) + C(r_j + 0)}{2}\right].$$

Mais nous sommes dans le cas où $A - a = 0$ et où par suite

$$\sum_j s_j'[C(r_j) - C(r_j - 0)] = \sum_j s_j''[C(r_j + 0) - C(r_j)] = 0.$$

Dès lors, on voit qu'on aura

$$\tfrac{1}{2}(A + a) = \sum_j s_j'[L - C(r_j)] + \sum_j s_j''[L - C(r_j)] = \sum_j s_j[L - C(r_j)],$$

en appelant s_j le saut total $s_j' + s_j''$ en r_j; d'où finalement

$$\int_V S(x)dC(x) = (L - l)(H - h) - \sum_j s_j[C(r_j) - l].$$

On remarquera que si l'on pose $\gamma(x) = C(x) - l$ pour avoir en $\gamma(x)$, comme en $S(x)$, une fonction (monotone bornée) ayant comme borne inférieure zéro et si l'on appelle Ω, Q les bornes supérieures de S et de $\gamma(x)$, on aura

$$\int_V S(x)d\gamma(x) = \Omega Q - \sum_j s_j\gamma(r_j) = \int_V S(x)dC(x).$$

Dans les cas où $\int_V f(x)dC(x)$ existe au moins suivant la définition I ou la définition II, l'intégrale $\int_V C(x)df(x)$ existe aussi, la condition à cet effet faisant intervenir symétriquement f et C. Or on a

$$\sum_{i=2}^{n} f(x_i)[C(x_i) - C(x_{i-1})] + \sum_{j=2}^{n} C(x_{j-1})[f(x_j) - f(x_{j-1})] = f(x_n)C(x_n) - f(x_1)C(x_1),$$

et

$$\sum_i f(x_i)[C(x_i) - C(x_{i-1})] + \sum_j C(x_{j-1})[f(x_j) - f(x_{j-1})] = LH - lh.$$

En prenant des divisions où chaque point de discontinuité commun à f et C s'introduise pour δ assez petit, et faisant tendre δ vers zéro, on obtient ainsi *la formule d'intégration par parties*

$$\int_\alpha^\beta f(x)dC(x) = [f(x)C(x)]_\alpha^\beta - \int_\alpha^\beta C(x)df(x),$$

et plus généralement

$$\int_V f(x)dC(x) = LH - lh - \int_V C(x)df(x),$$

formules établies ainsi dans tous les cas où $f(x)$ et $C(x)$ sont deux fonctions monotones et bornées telles que de chaque côté de chaque point de x du segment limité ou illimité V, l'une ou l'autre des fonctions $f(x)$, $C(x)$ soit continue.

Cas où f et C sont des fonctions à variations bornées. On peut étendre les résultats précédents aux intégrales où F et $v(x)$ sont à variations bornées sur V. On sait qu'alors $F(x)$ et $v(x)$ sont chacun différence de deux fonctions non-décroissantes $F = f_1 - f_2$, $v = C_1 - C_2$. Et si les variations totales de F et de v sur l'ensemble des points de V sont finies, les fonctions f_1, f_2, C_1, C_2 seront bornées sur V. Alors, la somme

$$\sigma = \sum_i F(\xi_i)\,[v(x_i) - v(x_{i-1})]$$

sera la somme algébrique de quatre sommes analogues mais formées chacune à partir de deux fonctions monotones bornées. Or les quatre fonctions f_1, f_2, C_1, C_2 peuvent être choisies de façon à n'avoir pas d'autres points de discontinuité à droite que F et v respectivement et n'avoir pas d'autres points de discontinuité à gauche que les mêmes fonctions. Dès lors: si de chaque côté de chaque point x, l'une au moins des fonctions F et v est continue, la somme σ tendra, quand $\delta \to 0$, vers une limite indépendante du choix des ξ_i dans les (x_{i-1}, x_i) et du choix de la suite des divisions D pourvu que tout point de discontinuité commun à F et v appartienne à la suite des D à partir d'une valeur assez petite de δ. Cette limite unique pourra être prise comme définition de $\int_V F(x)dv(x)$ quand F et v sont à variations totales bornées sur l'ensemble des points de V.

Seconde partie. Cas où dans $\int_V \varphi(x)dC(x)$, $\varphi(x)$ est continu mais le domaine d'intégration illimité

Dans ce qui précède, nous avons admis comme évident que l'extension de la définition et des propriétés classiques de $\int_\alpha^\beta \varphi(x)dC(x)$ où φ est continu et C monotone au cas où le domaine V d'intégration est illimité (dans un ou deux sens) ne présente pas de difficultés quand φ et C sont bornés. Dans ce cas, comme φ reste uniformément continu sur l'ensemble total des points du domaine d'intégration, on démontre, en effet, que la somme habituelle

$$\sum_i \varphi(\xi_i)\,[C(x_i) - C(x_{i-1})],$$

qui est ici une série absolument convergente, a une limite indépendante du choix des ξ_i dans les (x_{i-1}, x_i) et du choix de la suite des divisions D pourvu que δ tende vers zéro.

Nous allons examiner le cas moins simple où φ encore supposé continu n'est plus supposé borné sur le domaine illimité V pour lequel, pour préciser, nous prendrons $(-\infty, +\infty)$.

Une seconde définition de l'intégrale $\int_{-\infty}^{\infty} \varphi(x)dC(x)$. On peut définir l'intégrale $\int_{-\infty}^{\infty} \varphi(x)dC(x)$ comme la limite de $\int_a^b \varphi(x)dC(x)$ lorsque a et b tendent vers $-\infty$ et $+\infty$. Il est souvent plus commode d'utiliser une définition équivalente plus directe. Nous allons indiquer celle-ci.

Supposons donc que $C(x)$ soit une fonction monotone (par exemple, non-décroissante) et que $\varphi(x)$ soit une fonction continue quelconque. Pour assurer la validité de notre raisonnement, il nous sera nécessaire de supposer $C(x)$ bornée (ce qui n'est pas indispensable pour que l'intégrale ait un sens, mais qui est en tout cas réalisé dans le cas important où $C(x)$ est une "fonction des probabilités totales"). Soit maintenant $\cdots, x_{-m}, \cdots, x_{-1}, x_0, x_1, x_2, \cdots, x_n, \cdots$, une suite croissant de $-\infty$ à $+\infty$ et provisoirement arbitraire. On aura

$$J = \int_{-\infty}^{\infty} \varphi(x)dC(x) = \sum_{i=-\infty}^{\infty} \int_{x_{i-1}}^{x_i} \varphi(x)dC(x)\,,$$

si on suppose que l'intégrale du premier membre existe, au premier sens indiqué c'est à dire qu'elle est la limite de $\int_a^b \varphi(x)dC(x)$, de quelque façon que a et b tendent indépendamment vers $-\infty$ et $+\infty$. Cette dernière circonstance sera sûrement réalisée si l'on suppose que $\int_a^b |\varphi(x)|\,dC(x)$ est bornée quels que soient a et b, c'est à dire que l'intégrale J est absolument convergente, hypothèse où nous allons maintenant nous placer.

Or, en désignant par ξ_i un point arbitraire de l'intervalle et par m_i, M_i les bornes de $\varphi(x)$ dans cet intervalle, il est clair que $\int_{x_{i-1}}^{x_i} \varphi(x)dC(x)$ et $\varphi(\xi_i)[C(x_i) - C(x_{i-1})]$ sont compris entre $m_i[C(x_i) - C(x_{i-1})]$ et $M_i[C(x_i) - C(x_{i-1})]$. En posant $\omega_i = M_i - m_i$ on aura donc

$$\left| \int_{x_{i-1}}^{x_i} \varphi(x)dC(x) - \varphi(\xi_i)\,[C(x_i) - C(x_{i-1})] \right| \leqq \omega_i\,[C(x_i) - C(x_{i-1})]\,,$$

d'où

$$(1) \quad \left| \int_{x_{-j}}^{x_k} \varphi(x)dC(x) - \sum_{i=1-j}^{k} \varphi(\xi_i)\,[C(x_i) - C(x_{i-1})] \right| \leqq \sum_{i=1-j}^{k} \omega_i\,[C(x_i) - C(x_{i-1})]\,.$$

La fonction $\varphi(x)$ étant partout continue, il est toujours possible de choisir les x_i de sorte que les oscillations ω_i soient inférieures à un même nombre arbitraire

$\omega > 0$. Supposons qu'il en soit ainsi; on aura, en représentant les deux termes de (1) par $J_{j,k}$ et $S_{j,k}$

$$(2) \qquad | J_{j,k} - S_{j,k} | < \omega \sum_{i=1-j}^{k} [C(x_i) - C(x_{i-1})] = \omega(B - A) ,$$

A et B étant les deux bornes de $C(x)$ sur toute l'ensemble des valeurs de x. Appelons $J'_{j,k}$, $S'_{j,k}$, ω'_i, J', ω_i les valeurs prises par $J_{j,k}$, $S_{j,k}$, ω_i, J, ω quand on y remplace φ par $|\varphi|$. On a nécessairement $\omega'_i \leqq \omega_i \leqq \omega$. Donc les ω'_i ont bien une borne supérieure finie $\omega' \leqq \omega$. On aura de même

$$S'_{j,k} < J'_{j,k} + \omega'(B - A) \leqq \int_{-\infty}^{\infty} | \varphi(x) | \, dC(x) + \omega'(B - A) .$$

Quand j et k croissent, $S'_{j,k}$ croît, ou du moins ne décroît pas, et reste inférieur, d'après cette inégalité, à un nombre indépendant de j et de k. Donc la série

$$(3) \qquad S = \sum_{i=-\infty}^{\infty} \varphi(\xi_i) \, [C(x_i) - C(x_{i-1})]$$

est absolument convergente et l'on a

$$S = \lim_{\substack{j \to -\infty \\ k \to +\infty}} S_{j,k} .$$

On tire alors de l'inégalité (2), en passant à la limite, $| J - S | < \omega(B - A)$, d'où $J = \lim\limits_{\omega \to 0} S$. Réciproquement, lorsque les x_i forment une suite croissant de $-\infty$ à $+\infty$ et choisis de façon que les oscillations ω_i de $\varphi(x)$ dans les intervalles x_{i-1}, x_i aient une borne supérieure finie ω, si la série S est absolument convergente pour au moins un choix des ξ_i dans les intervalles x_{i-1}, x_i, alors l'intégrale J est absolument convergente et on a $J = \lim\limits_{\omega \to 0} S$. En effet, on a encore

$$| S'_{j,k} - J'_{j,k} | \leqq \omega'(B - A) ,$$

d'où

$$J'_{j,k} \leqq S'_{j,k} + \omega'(B - A) \leqq \sum_{i=-\infty}^{\infty} | \varphi(\xi) | \, [C(x_i) - C(x_{i-1})] + \omega'(B - A) .$$

Donc $J'_{j,k}$ ayant une borne supérieure finie indépendante de j et de k, l'intégrale J est absolument convergente. Alors en faisant tendre j et k vers $-\infty$ et $+\infty$ dans la formule $| S_{j,k} - J_{j,k} | \leqq \omega(B - A)$, on aura $| S - J | \leqq \omega(B - A)$, d'où $J = \lim\limits_{\omega \to \infty} S$.

Remarque. Nous venons de démontrer que si $\varphi(x)$ est une fonction partout continue, si $C(x)$ est une fonction monotone bornée, la condition pour que l'intégrale $J = \int_{-\infty}^{\infty} \varphi(x) dC(x)$ soit absolument convergente est qu'il existe au moins un nombre $\omega > 0$, et, — en divisant la droite illimitée par une suite crois-

sante de nombres x_i tels que l'oscillation de $\varphi(x)$ dans chaque intervalle (x_{i-1}, x_i) soit $< \omega -$, un choix de nombres ξ_i dans les intervalles respectifs (x_{i-1}, x_i) tels que la série (3) soit absolument convergente.

Et s'il en est ainsi, on a $J = \lim_{\omega \to 0} S$ quel que soit le choix des ξ_i dans les (x_{i-1}, x_i). La démonstration suppose essentiellement que $C(x)$ est borné. On peut voir de plus que la validité de l'énoncé cesse si l'on supprime cette condition. Il suffit de prendre l'exemple suivant.

Considérons le cas particulier où $C(x) = x$ et où $\varphi(x)$ est une fonction continue paire, jamais négative, et choisie dans chaque intervalle $2(k - 1) \leqq x \leqq 2k$, telle que l'on y ait $0 \leqq \varphi(x) \leqq 1/k$, d'où il résulte $0 \leqq \int_{2k-2}^{2k} \varphi(x)dx \leqq 2k^{-1}$, et que de plus l'on ait $\varphi(2k - 1) = k^{-1}$, $\int_{2k-2}^{2k} \varphi(x)dx = 2k^{-2}$, pour $k > 1$. Alors, on voit qu'ici l'intégrale $J = \int_{-\infty}^{\infty} \varphi(x)dC(x)$ sera finie et égale à $4 \sum_{k=1}^{\infty} k^{-2}$. Pourtant, soit ω un nombre positif arbitraire; il existe un entier N tel que $2 < N\omega$. On pourra donc prendre les x_i tels qu'à partir d'un certain rang q (variable avec ω) on ait $x_q = 2N$; $x_{q+1} = 2(N + 1), \cdots , x_{q+p} = 2(N + p), \cdots$, car alors dans (x_{q+p-1}, x_{q+p}), on aura $| \varphi(x') - \varphi(x'') | < \varphi(x') + \varphi(x'') < 2/(N + p) < \omega$. Alors en prenant $\xi_{q+p} = 2(N + p) - 1$, on aura pour tout entier K

$$S > 2 \sum_{k=q}^{K+q} \varphi(\xi_k)(x_k - x_{k-1}) > \sum_{p=0}^{K} \frac{1}{2(N + p) - 1},$$

et le dernier terme croît indéfiniment avec K. Donc J est fini et S est infini, et infini quel que soit ω; on ne peut donc avoir $J = \lim_{\omega \to 0} S$.

La proposition s'étend d'ailleurs au cas où $C(x)$ est remplacée par une fonction $v(x)$ à variation bornée en spécifiant convenablement les conditions d'application.

On sait que si $V(x)$ est la variation totale de $v(x)$ dans l'intervalle (a, x), $(a < x)$ et si l'on pose $v(x) = v(a) + P_1(x) - N(x)$, $V(x) = P_1(x) + N(x)$, les fonctions $P_1(x)$ et $N(x)$ sont non-décroissantes de sorte qu'en posant $P(x) = v(a) + P_1(x)$, $v(x)$ est la différence de deux fonctions non-décroissantes, $v(x) = P(x) - N(x)$.

Le résultat subsiste pour $x < a$, si l'on égale alors à $-V(x)$ la variation totale de $v(x)$ de x à a. Si donc $v(x)$ est à variation totale bornée *dans tout intervalle fini*, $v(x)$ est la différence de deux fonctions *partout* non-décroissantes. Si maintenant la variation totale de $v(x)$ dans un intervalle (a, b) est bornée quand b tend vers $+\infty$ comme on a $| v(x) - v(a) | \leqq V(x)$, $v(x)$ restera borné comme $V(x)$ pour $x > a$ et il en sera de même de $P(x)$ et de $N(x)$; en particulier, $P(+\infty)$ et $N(+\infty)$ auront une signification et des valeurs déterminées. Si, même, la variation totale de $v(x)$ dans tout intervalle a une borne supérieure indépendante de cet intervalle, les fonctions $P(x)$, $N(x)$ seront bornées supérieurement et inférieurement. C'est dans ce cas que nous pourrons généraliser le théorème établi ci-dessus, en posant:

$$\int_a^b \varphi(x)dv(x) = \int_a^b \varphi(x)dP(x) - \int_a^b \varphi(x)dN(x).$$

La condition nécessaire et suffisante pour que les deux intégrales du second membre soient absolument convergentes est évidemment que $\int_{-\infty}^{\infty} |\varphi(x)|\, d[P(x) + N(x)]$ et par suite $\int_{-\infty}^{\infty} |\varphi(x)|\, dV(x)$ soit convergente. On a d'ailleurs

$$\left| \int_{-\infty}^{\infty} \varphi(x) dv(x) \right| \leqq \int_{-\infty}^{\infty} |\varphi(x)|\, dV(x).$$

On a alors la proposition suivante. Si $\varphi(x)$ est une fonction continue partout, si $v(x)$ est une fonction dont la variation totale dans un intervalle a, b a une borne supérieure indépendante de a, b, alors pour que les intégrales $\int_{-\infty}^{\infty} \varphi(x) dv(x)$ et $\int_{-\infty}^{\infty} |\varphi(x)|\, dV(x)$ soient à la fois convergentes, il faut et il suffit que, pour au moins une suite de nombres x_i croissant de $-\infty$ à $+\infty$ et tels que les oscillations ω_i de $\varphi(x)$ dans les intervalles (x_{i-1}, x_i) aient une borne supérieure finie ω, et pour au moins une suite de nombres pris dans les intervalles (x_{i-1}, x_i), la série

$$\sum_{i=-\infty}^{\infty} |\varphi(\xi_i)|\, [V(x_i) - V(x_{i-1})]$$

soit convergente. Et alors on aura

$$\int_{-\infty}^{\infty} \varphi(x) dv(x) = \lim_{\omega \to 0} \sum_{i=-\infty}^{\infty} \varphi(\xi_i)\, [v(x_i) - v(x_{i-1})]$$

quel que soit le choix des ξ_i dans les (x_{i-1}, x_i).

Observons que si $\varphi(x)$ est uniformément continue non seulement dans tout intervalle fini, mais sur la droite illimitée (ce qui, par exemple, a lieu pour $\varphi(x) = x$, mais non pour $\varphi(x) = x^2$) la condition que tous les ω_i soient inférieurs à ω sera réalisée en prenant tous les $x_i - x_{i-1}$ inférieurs à un même nombre δ assez petit. Alors l'égalité (2) subsistera en remplaçant ω par δ. Si $\varphi(x)$ est continue et bornée, $\varphi(x)$ est uniformément continue et le résultat précédent s'applique. Mais ce n'est pas le seul cas. En particulier, pour que $\int_{-\infty}^{\infty} x dC(x)$ existe, il faut et il suffit qu'il existe au moins *une* suite de nombres x_i croissant de $-\infty$ à $+\infty$, dont les intervalles $x_i - x_{i-1}$ ont une borne supérieure finie δ et *une* suite de nombres ξ_i dans les intervalles (x_{i-1}, x_i), telles que la série

$$S = \sum_{i=-\infty}^{\infty} \xi_i\, [C(x_i) - C(x_{i-1})]$$

soit *absolument* convergente. Et alors, on a $\int_{-\infty}^{\infty} x dC(x) = \lim_{\delta \to 0} S$. Ce cas particulier est intéressant en Calcul des Probabilités où la valeur moyenne d'une variable aléatoire s'exprime précisément sous la forme $\int_{-\infty}^{\infty} x dC(x)$.

UNIVERSITÉ DE PARIS.

NEW THEOREMS AND METHODS IN DETERMINANT THEORY

By Leonard M. Blumenthal

Introduction. If to each ordered pair of undefined elements p, q of an abstract space (set) S, a real, non-negative number pq can be attached such that $pq = qp$, and $pq = 0$ if and only if p is identical with q, the space S is said to be *semimetric*. The elements p, q may be spoken of as points of the space, with pq as their distance. A given semimetric space S is *characterized metrically* when conditions are stated (in terms of distance relations) which are necessary and sufficient for any semimetric space satisfying them to be mapped isometrically (congruently) upon S. Among those semimetric spaces which have been characterized metrically are the n-dimensional euclidean,[1] spherical,[2] and hyperbolic[3] spaces.

In this paper results obtained in the metric characterization of these spaces are introduced for the purpose of deriving new theorems concerning certain types of symmetric determinants. The application of isometric geometry to determinant theory furnishes a new and powerful impetus for its development. By such an application one obtains elegant proofs of novel and interesting theorems. These new methods are well adapted (1) for proving whole chains of theorems, as in §§1 and 5, (2) for the determination of relations between the elements of a determinant, as in Theorems 3.1 and 5.2, and (3) for ascertaining the sign of certain determinants, whose elements are not explicitly known.

While only determinants with real elements are treated in this paper, the development of the theory of complex metric spaces, already under way, may be expected to furnish results that can be applied to determinants with complex elements.[4]

Received August 29, 1935; presented to the American Mathematical Society, October 26, 1935, under the title *Isometric geometry methods in determinant theory*. The author is National Research Fellow.

[1] Menger, *Untersuchungen über allgemeine Metrik*, Mathematische Annalen, vol. 100 (1928), pp. 75–163. This paper is divided into three parts; the second part (Zweite Untersuchung, pp. 113–141) contains a characterization of the n-dimensional euclidean space in terms of relations between the distances of its points. In a later paper, *New foundation of euclidean geometry*, American Journal of Mathematics, vol. 53 (1931), pp. 721–745, the concept of quasi-congruence order is introduced.

[2] Blumenthal, *Concerning spherical spaces*, American Journal of Mathematics, vol. 57 (1935), pp. 51–61. See also vol. 55 (1933), pp. 619–640, as well as L. Klanfer, *Metrische Charakterisierung der Kugel*, Ergebnisse eines mathematischen Kolloquiums, Wien, Heft 4 (1933), pp. 43–45.

[3] Blumenthal, *The metric characterization of the n-dimensional hyperbolic space*, Bull. Amer. Math. Soc., vol. 41 (1935), p. 485 (Abstract).

[4] A. Wald, *Komplexe und indefinite Räume*, Ergebnisse eines mathematischen Kolloquiums, Wien, Heft 5 (1933), pp. 32–42.

Part I. The determinant $| r_{ij} |, r_{ij} = r_{ji}, r_{ii} = 1 \; (i, j = 1, 2, \cdots, m)$

1. In this section we deal with determinants of the above form which satisfy the additional hypothesis that $r_{ij} > 1 \; (i, j = 1, 2, \cdots, m; i \neq j)$. Since these determinants play an important rôle in the metric characterization of the n-dimensional hyperbolic space, we shall denote them by the letter H.

We prove first the following

THEOREM 1.1. *Let $p_1, p_2, \cdots, p_{n+1}$ be $n + 1$ points of the n-dimensional hyperbolic space $H_{n,r}$, of curvature $-1/r^2$. The determinant $| \cosh p_ip_j/r |$ $(i, j = 1, 2, \cdots, n + 1)$ has the sign $(-1)^n$ if $p_1, p_2, \cdots, p_{n+1}$ are independent (i.e., not in $H_{m,r}$, $m < n$), and vanishes otherwise.*

If we multiply the first column of the determinant by $\cosh p_1p_k/r$, and subtract the result from the k-th column, $k = 2, 3, \cdots, n + 1$, we find, after applying the law of cosines for hyperbolic geometry,

$$\left| \cosh \frac{p_ip_j}{r} \right| = (-1)^n \cdot | \cos p_1 : p_ip_j |_{(i,j=2,3,\cdots,n+1)} \cdot \prod_{k=2}^{n+1} \sinh^2 \frac{p_1p_k}{r},$$

where $p_1 : p_ip_j$ denotes the angle formed at p_1 by the rays p_1p_i, p_1p_j, and $| \cos p_1 : p_ip_j | \; (i, j = 2, 3, \cdots, n + 1)$ denotes the determinant of order n of these elements. It remains to show that this determinant is positive if $p_1, p_2, \cdots, p_{n+1}$ are independent points, and zero otherwise.

But this is immediate, for the "n-bein" formed in $H_{n,r}$ by the rays joining the points $p_2, p_3, \cdots, p_{n+1}$ to p_1 may, as is well known, be imbedded *isogonally* in a euclidean n-dimensional space. Since, now, the determinant $| \cos p_1 : p_ip_j |$ $(i, j = 2, 3, \cdots, n + 1)$ is invariant under an isogonal transformation, and since this function formed for n rays in a euclidean space is positive if the rays are independent and zero otherwise, the lemma is proved.[5]

THEOREM 1.2. *If the determinant H is of order $m > n + 3$, and if* (i) *for each integer $1 \leqq k \leqq n$ each principal minor of order $k + 1$ of H vanishes or has the sign $(-1)^k$, and* (ii) *each principal minor of order $n + 2$ vanishes, then H vanishes.*

Proof. Since $r_{ij} > 1 \; (i \neq j)$, we may write

$$r_{ij} = \cosh p_ip_j/r \qquad (i, j = 1, 2, \cdots, m),$$

where the points $p_1, \cdots, p_m$ form a semimetric set and $r > 0$. From the hypotheses (i), (ii) we may conclude that each set of $n + 2$ of the points

[5] The relation satisfied by the ten distances of five points in a three-dimensional hyperbolic space was first given by Schering, *Die Schwerkraft im Gaussischen Räume*, Göttinger Nachrichten, 1870, pp. 311–321, without, however, indicating how it was obtained. In a later paper (Göttinger Nachrichten, 1873, pp. 13–21) he stated the analogous relation for $n + 2$ points in an n-dimensional hyperbolic space. P. Mansion, Ann. Soc. Sc. Brux., vol. 15 (1890–1891), pp. 8–11; vol. 19 (1894–1895), pp. 189–193 deduces Schering's five-points relation. The method used in the lemma above is completely different from that used by Mansion. So far as the writer is aware, the determination of the sign of the function $| \cosh p_ip_j/r |$ for $n + 1$ independent points of $H_{n,r}$, given above, is not in the earlier literature.

$p_1, \cdots, p_m$ is congruent with $n+2$ points of the hyperbolic n-dimensional space H_n of curvature $-1/r^2$, while from the fact that H_n has the quasi-congruence order $n+2$, it follows that the m points are congruent with m points of H_n.[6]

Using Theorem 1.1, we conclude that $H=0$.

2. Let us suppose, now, that $0 \leqq r_{ij} < 1$ $(i, j = 1, 2, \cdots, m;\ i \neq j)$. Determinants of the form we are considering in Part I which satisfy this additional requirement we denote by Δ_p. Concerning determinants Δ_p we prove the following

THEOREM 2.1. *If the determinant Δ_p is of order $m > 4$, and has each third-order principal minor equal to zero, the determinant Δ_p vanishes.*

Proof. From the hypotheses made on the elements r_{ij} we may set $r_{ij} = \cos p_ip_j/r$, $r > 0$ $(i, j = 1, 2, \cdots, m)$, with the points $p_1, p_2, \cdots, p_m$ forming a semimetric set, and $p_ip_j \leqq \pi r/2$ for every pair of indices.

From the metric characterization of the circle[7] we may conclude from the vanishing of every third-order principal minor, together with $p_ip_j < \pi r$ $(i, j = 1, 2, \cdots, m)$, that each triple of points contained in the m points $p_1, p_2, \cdots, p_m$ is congruent with three points of a circle of radius r. (The distance of two points of the circle is defined as the length of the shorter arc joining them.) Each triple is, moreover, linear (i.e., congruent with three points of a line), for otherwise the sum of the three distances would equal $2\pi r$, which is impossible.

Since the line has the quasi-congruence order[8] 3, and by hypothesis $m > 4$, it follows that the m points $p_1, p_2, \cdots, p_m$ are linear and consequently congruent with m points of a circle of radius r (since no distance exceeds $\pi r/2$). Then the determinant Δ_p vanishes.

It is believed that Theorem 2.1 is the first link in a chain of theorems which we state as follows:

If the determinant Δ_p of order $m > n+3$ (n an integer) is such that (i) *every principal minor of order not exceeding $n+1$ is positive or zero,* (ii) *every principal minor of order $n+2$ is zero, the determinant vanishes.*

3. In this section we deal with determinants

$$\Delta_N = |\, r_{ij} \,|, \quad r_{ij} = r_{ji}, \ -1 < r_{ij} \leqq 0 \ (i \neq j), \quad r_{ii} = 1, \quad (i, j = 1, 2, \cdots, m),$$

and concerning them we prove the following

THEOREM 3.1. *If the determinant Δ_N is of order $m > 4$, and all of its third-order principal minors vanish, then $r_{ij} = -\frac{1}{2}$ $(i, j = 1, 2, \cdots, m;\ i \neq j)$ and $\Delta_N = -\frac{1}{2}(3/2)^{m-1}(m-3)$.*

[6] Blumenthal; see footnote 3.

[7] Blumenthal and Garrett, *Characterization of spherical and pseudo-spherical sets of points*, American Journal of Mathematics, vol. 55 (1933), pp. 619–640. See §I, p. 620, for the circle.

[8] Menger, *New foundation of euclidean geometry*, p. 727.

Proof. We write $r_{ij} = \cos p_ip_j/r$, $r > 0$, $(i, j = 1, 2, \cdots, m)$, where the m points $p_1, p_2, \cdots, p_m$ form a semimetric set. Since $-1 < r_{ij} \leqq 0$, we may assume that $\pi r/2 \leqq p_ip_j < \pi r$ for distinct values of the indices. From the vanishing of each third-order principal minor of Δ_N, together with $p_ip_j < \pi r$, it follows that the set of points $p_1, p_2, \cdots, p_m$ has all of its triples d-cyclic, with $d = \pi r$. But *no one of these triples is linear*, for in a linear triple the sum of two of the distances equals the third, while any triple contained in $p_1, p_2, \cdots, p_m$ has the sum of two of its distances greater than or equal to πr, while the third distance is less than πr. It follows readily that the set $p_1, p_2, \cdots, p_m$ forms a *proper* pseudo d-cyclic set and hence is equilateral, with each distance equal[9] to $2d/3$. Hence $p_ip_j = 2d/3$, and $r_{ij} = \cos 2d/3r = -\frac{1}{2}$ $(i, j = 1, 2, \cdots, m; i \neq j)$. Such a determinant is easily evaluated to yield the value $-\frac{1}{2}(3/2)^{m-1}(m-3)$.

Theorems 2.1 and 3.1 are in striking contrast. In the latter theorem the hypotheses serve to *fix every element* in the determinant.

There is reason to believe that the following chain of propositions is valid, but its proof must await, it seems, further development of the theory of pseudo r-spheric $(S_{n,r})$ sets of points.

If the determinant Δ_N is of order $m > n + 3$ (n an integer) and is such that (i) *every principal minor of order less than or equal to $n + 1$ is positive and* (ii) *every principal minor of order $n + 2$ is zero, then*

$$r_{ij} = -1/(n+1) \qquad (i, j = 1, 2, \cdots, m; i \neq j),$$

and

$$\Delta_N = -\frac{1}{n+1}\left(\frac{n+2}{n+1}\right)^{m-1}(m - n - 2).$$

It is interesting to compare Theorems 2.1 and 3.1 with another theorem which the writer has proved in an earlier paper, and which is, in a sense, the "union" of these two theorems.[10] Let us denote by $\Delta_N(-\frac{1}{2})$ the determinant r_{ij}, with $r_{ij} = -\frac{1}{2}$ $(i \neq j)$, $r_{ii} = 1$, $(i, j = 1, 2, \cdots, m)$.

THEOREM. *If the determinant $\Delta = |r_{ij}|$, $r_{ij} = r_{ji}$, $-1 < r_{ij} < 1$ $(i \neq j)$, $r_{ii} = 1$, $(i, j = 1, 2, \cdots, m)$, is of order greater than 4, and has each of its third-order principal minors equal to zero, either Δ vanishes or $\Delta = \Delta_N(-\frac{1}{2})$.*

It is noted that increasing the range of the elements r_{ij} $(i \neq j)$ from the interval $(-1, 0)$, open at the left, as in Theorem 3.1, to the open interval $(-1, 1)$, as in the theorem above, does not sensibly increase the class of *non-vanishing*

[9] Blumenthal, *A complete characterization of proper pseudo d-cyclic sets of points*, American Journal of Mathematics, vol. 54 (1932), pp. 387–396. A *proper* pseudo d-cyclic set is, by definition, a pseudo d-cyclic set that contains neither a convex tripod nor a pseudo-linear quadruple. A convex tripod has three of its triples linear, while all four triples of a pseudo-linear quadruple are linear. Hence, a pseudo d-cyclic set that contains no linear triples cannot contain a convex tripod or a pseudo-linear quadruple, and is, therefore, proper.

[10] Blumenthal, *A chain of determinant theorems arising from the characterization of pseudo r-spheric $(S_{n,r})$ sets*, American Journal of Mathematics, vol. 56 (1934), pp. 225–232.

determinants that satisfy the remaining hypotheses of these two theorems. In the first case we have proved that such determinants have each element outside of the principal diagonal equal to $-\frac{1}{2}$; while the earlier theorem proves that each element outside the principal diagonal is equal to either $\frac{1}{2}$ or $-\frac{1}{2}$, and that, further, the signs of the elements are so distributed that by multiplying certain rows and the corresponding columns by -1, the determinant is made identical with $\Delta_N(-\frac{1}{2})$.

Though Theorem 2.1 is implied by the theorem quoted above, its proof demands far less than what is given by this theorem. This reason, together with the desirability of developing methods that can be generalized to prove the chain of propositions of which Theorem 2.1 is the first link, has led us to give a short *independent* proof of it.

Part II. The determinant[11] $D = |r_{ij}|, r_{ij} = r_{ji} > 0\ (i \neq j), r_{ii} = 0,$

$$r_{0i} = 1\ (i \neq 0)\ , \qquad (i, j = 0, 1, \cdots, m)\ .$$

4. It is convenient to emphasize the bordering of this determinant by writing it symbolically in the form $D = \begin{vmatrix} 0 & 1 \\ 1 & r_{ij} \end{vmatrix}$ $(i, j = 1, 2, \cdots, m)$. We give first a theorem concerning fifth-order determinants of this type.

THEOREM 4.1. *If the determinant* $D = \begin{vmatrix} 0 & 1 \\ 1 & r_{ij} \end{vmatrix}$ $(i, j = 1, 2, 3, 4)$ *has each of its four bordered fourth-order principal minors negative or zero, then, for* $0 \leqq k \leqq \frac{1}{2}$, *the determinant* $D^{(k)} = \begin{vmatrix} 0 & 1 \\ 1 & r_{ij}^{k} \end{vmatrix}$ $(i, j = 1, 2, 3, 4)$, *of non-negative elements, has* (1) *each of its four bordered fourth-order principal minors negative,* (2) $D^{(k)} > 0$, $0 \leqq k < \frac{1}{2}$, $D^{(\frac{1}{2})} \geqq 0$, *and* (3) $k = \frac{1}{2}$ *is the greatest exponent for which both* (1) *and* $D^{(k)} \geqq 0$ *are valid.*

Proof. Since $r_{ij} = r_{ji} > 0\ (i \neq j)$, $r_{ii} = 0$, we may set $r_{ij} = p_ip_j^2$, $(i, j = 1, \cdots, 4)$, where p_1, p_2, p_3, p_4 form a semimetric set. By hypothesis, each bordered fourth-order principal minor of D is negative or zero. This is a

[11] For the cases $m = 3, 4, 5$ this bordered determinant has a long history. Early papers by Cayley [Cambridge Mathematical Journal, vol. 2 (1841), p. 268] and Sylvester show that if five points p_i $(i = 1, 2, \cdots, 5)$ are in a three-dimensional euclidean space, the sixth-order determinant D, in which r_{ij} $(i, j = 1, 2, \cdots, 5)$ is the square of the distance p_ip_j, vanishes. This result had been obtained earlier in a different form by Lagrange, Carnot, and others. Since Cayley was among the first to discuss the behavior of this determinant when it was formed for the points of a euclidean space, and since Menger (loc. cit.) was the first to characterize metrically euclidean space by expressing the (necessary and sufficient) distance relations of semimetric spaces congruent with it in terms of the sign of the determinant D formed for these distances, the determinant may well be called the Cayley-Menger determinant.

necessary and sufficient condition that each triple of points contained in the four points be congruent with three points of a euclidean[12] plane R_2.

To prove part (1) of the theorem, we introduce a semimetric set of four points, p_1', p_2', p_3', p_4', with $r_{ij}^k = (p_i'p_j')^2$ $(i, j = 1, \cdots, 4)$, and show that for $0 \leqq k \leqq \frac{1}{2}$, each triple of "primed" points is congruent with a triple of the plane R_2, and *is not linear*. Consider the triple p_1', p_2', p_3'. Now it has been shown[13] that if $f(x)$ is any monotonic increasing function of the real variable x, which is, further, a concave function of this variable, and vanishes for $x = 0$, and if $\bar{x}_1, \bar{x}_2, \bar{x}_3$ are any three positive values satisfying the triangle inequality, the positive numbers $f(\bar{x}_1), f(\bar{x}_2), f(\bar{x}_3)$ satisfy the *strict* triangle inequality.

Since the positive branch of the function x^α, $0 < \alpha < 1$, $x \geqq 0$, is evidently a function of this type, and since p_1p_2, p_2p_3, p_1p_3 are positive numbers and satisfy the triangle inequality, we see that the numbers $(p_1p_2)^k$, $(p_2p_3)^k$, $(p_1p_3)^k$ satisfy the *strict* triangle inequality (and hence the points p_1', p_2', p_3' are non-linear and congruent with three points of a plane) for $0 \leqq k \leqq \frac{1}{2}$. Hence the bordered fourth-order principal minor $\begin{vmatrix} 0 & 1 \\ 1 & r_{ij}^k \end{vmatrix}$ $(i, j = 1, 2, 3)$ is negative. Similarly, the other bordered fourth-order principal minors are shown to be negative, and part (1) of the theorem is proved.

Now, since each triple contained in the four points p_i' $(i = 1, 2, 3, 4)$ is congruent with a planar triple, the four points determine twelve angles, namely, the angles of the four planar triangles. We remark that *each of these twelve angles is acute, for* $0 \leqq k < \frac{1}{2}$. For, let $(p_{i_1}'p_{i_2}'p_{i_3}')$ be any one of these angles. We have

$$\cos(p_{i_1}'p_{i_2}'p_{i_3}') = \frac{(p_{i_1}'p_{i_2}')^2 + (p_{i_2}'p_{i_3}')^2 - (p_{i_1}'p_{i_3}')^2}{2(p_{i_1}'p_{i_2}')(p_{i_2}'p_{i_3}')}$$

$$= \frac{(p_{i_1}p_{i_2})^{2k} + (p_{i_2}p_{i_3})^{2k} - (p_{i_1}p_{i_3})^{2k}}{2(p_{i_1}p_{i_2})^k \cdot (p_{i_2}p_{i_3})^k}$$

and since $0 \leqq k < \frac{1}{2}$, then $0 \leqq 2k < 1$, and it follows that $(p_{i_1}p_{i_2})^{2k} + (p_{i_2}p_{i_3})^{2k} > (p_{i_1}p_{i_3})^{2k}$. Hence $\cos(p_{i_1}'p_{i_2}'p_{i_3}') > 0$, and since $(p_{i_1}'p_{i_2}'p_{i_3}')$ is an angle of a triangle, we conclude that it is acute.

To prove part (2), we show first that $D^{(\frac{1}{2})} = \begin{vmatrix} 0 & 1 \\ 1 & r_{ij}^{\frac{1}{2}} \end{vmatrix}$ $(i, j = 1, 2, 3, 4)$ is positive or zero. To do this, we assume the contrary and deduce from this assumption a contradiction. The function $D^{(k)}$ is a continuous function of k,

[12] This is at once evident upon developing one of these minors. Replacing r_{ij} by $p_ip_j^2$, we find, for example, that the minor formed for the points p_1, p_2, p_3 can be written in the form $-(p_1p_2 + p_2p_3 + p_1p_3)(p_1p_2 + p_2p_3 - p_1p_3)(p_1p_2 - p_2p_3 + p_1p_3)(-p_1p_2 + p_2p_3 + p_1p_3)$, and since, by hypothesis, this minor is negative or zero, it follows that the numbers p_1p_2, p_2p_3, p_1p_3 satisfy the triangle inequality; that is, the points p_1, p_2, p_3 are congruent with three points of a plane.

[13] Blumenthal, *Note on the euclidean four-points property*, Ergebnisse eines mathematischen Kolloquiums, Wien, Heft 7 (in press).

which is positive for $k = 0$ and, by our assumption, negative for $k = \frac{1}{2}$. Hence, there exists at least one value k' such that $D^{(k')} = 0$, and $0 < k' < \frac{1}{2}$.

Consider, now, four points p_i^* $(i = 1, 2, 3, 4)$ forming a semimetric set and such that $(p_i^* p_j^*)^2 = r_{ij}^{k'} = (p_i p_j^2)^{k'}$. Then $p_i^* p_j^* = (p_i p_j)^{k'}$, and since $k' < 1$, it follows (1) that each triple contained in the four points p_i^* $(i = 1, 2, 3, 4)$ is congruent to a planar triple, and (2) since $k' < \frac{1}{2}$, each of the twelve angles determined by the "starred" points is *acute*. Combining $D^{(k')} = 0$ with (1), we conclude[14] that $p_1^*, p_2^*, p_3^*, p_4^*$ are congruent with four points of the plane R_2. This yields the contradiction sought, for of the twelve angles determined by four points of a plane, at least one angle is greater than or equal to a right angle. Hence, not all of the twelve angles determined by the points $p_1^*, p_2^*, p_3^*, p_4^*$ can be acute. This contradicts (2) above. We conclude, then, that $D^{(\frac{1}{2})} \geqq 0$.

In showing that $D^{(\frac{1}{2})} \geqq 0$ we have shown that $D^{(k)}$ cannot vanish for $0 < k < \frac{1}{2}$, and since $D^{(0)}$ is positive, it follows that $D^{(k)} > 0$ for $0 \leqq k < \frac{1}{2}$, and part (2) is proved.

We prove part (3) of the theorem by means of an example. Consider the six numbers $r_{13} = r_{24} = 4$, $r_{12} = r_{23} = r_{34} = r_{14} = 1$. It is easily seen that the determinant D formed for these values has each of its four bordered fourth-order principal minors equal to zero, and hence satisfies the hypothesis of the theorem. If, now, we form the determinant $D^{(k)}$ where $k = \frac{1}{2}(1 + \epsilon)$, $0 \leqq \epsilon < 1$, we find that each of the four bordered fourth-order principal minors of $D^{(k)}$ is negative, while $D^{(k)} = -32 \cdot 2^{2\epsilon} \cdot (2^\epsilon - 1)$. Hence, for $\epsilon > 0$ the determinant is negative; that is, for $k > \frac{1}{2}$, $D^{(k)} < 0$, and the proof is complete.

Several remarks may be made concerning the foregoing theorem. It is of interest to observe that though *no hypothesis is made concerning the sign of the determinant D*, the extraction of positive k-th roots of its elements is for $0 \leqq k < \frac{1}{2}$ sufficient to make the new determinant $D^{(k)}$ *positive*, while $D^{(\frac{1}{2})} \geqq 0$. This fact has an important geometrical application, which we give in the following

COROLLARY. *If p_i $(i = 1, 2, 3, 4)$ are any four points of a metric space, for any non-negative number k, not exceeding $\frac{1}{2}$, there exist four points p_i' $(i = 1, 2, 3, 4)$ of a euclidean three-dimensional space such that $p_i' p_j' = (p_i p_j)^k$ $(i, j = 1, 2, 3, 4)$.*

Thus, if M is any metric space, and we denote by $M^{(k)}$ the space derived from M by taking the positive k-th power of its metric, then, for $0 \leqq k \leqq \frac{1}{2}$, the space $M^{(k)}$ has every quadruple of its points congruent to four points of a euclidean space.

The question arises whether the property proved by Theorem 4.1 is peculiar to the fifth-order determinants of the type considered, or whether, when the range of k is kept fixed, the theorem may be extended to determinants of higher order, the obvious additional hypotheses on higher ordered principal minors being made. In the proof of Theorem 4.1 we made use of the fact, from elementary geometry, that not all of the twelve angles determined by four distinct points of a plane can be acute. This fact, sufficient for our proof, turns out to be also necessary. But for $n > 2$ it is *not true* that if $n + 2$ points are in an

[14] Menger, Mathematische Annalen, loc. cit., p. 136.

n-dimensional space R_n, not all of the $3\cdot\binom{n+2}{3}$ angles determined by the $n+2$ points can be acute.[15] Thus the theorem holds only for fifth-order determinants. Its restricted character, when regarded only as a theorem on determinants, is, however, atoned for when its interesting geometrical applications are taken into account. In addition, it raises two questions: (1) if a fifth-order determinant D satisfies the hypotheses of the theorem, what is the most general function f such that the determinant $f(D)$ satisfies the conclusions that every fourth-order bordered principal minor of $f(D)$ is negative *or zero*, and $f(D) \geqq 0$, where $f(D) = \begin{vmatrix} 0 & 1 \\ 1 & f(r_{ij}) \end{vmatrix}$ $(i, j = 1, 2, 3, 4)$, and (2) for what functions can the theorem be extended to higher ordered determinants?

We remark, finally, that if it is assumed that D is non-negative, the range of k for which the theorem with this added hypothesis is valid is probably the interval $(0, 1)$.

5. A chain of theorems concerning determinants $D = \begin{vmatrix} 0 & 1 \\ 1 & r_{ij} \end{vmatrix}$ $(i, j = 1, 2, \cdots, m)$ is easily obtained from the characterization of the n-dimensional euclidean space R_n.

THEOREM 5.1. *Let D be of order $m + 1 > n + 4$ (n an integer) and suppose* (i) *for every integer $k \leqq n + 1$, each bordered principal minor of order $k + 1$ has the sign $(-1)^k$ or vanishes,* (ii) *every bordered principal minor of order $n + 3$ vanishes; then the determinant D vanishes.*

Proof. We introduce a semimetric set of points p_i $(i = 1, 2, \cdots, m)$, $m > n + 3$, such that $p_ip_j^2 = r_{ij}$ $(i, j = 1, 2, \cdots, m)$. From hypothesis (i) it follows directly that every set of $n + 1$ points contained in the m points is congruent with $n + 1$ points of the R_n; this, combined with hypothesis (ii) justifies a similar remark[16] concerning each set of $n + 2$ points contained in p_i $(i = 1, 2, \cdots, m)$. But the R_n has been shown to have the quasi-congruence order $n + 2$; that is, in order that a semimetric set consisting of *more than* $n + 3$ points be congruent with a subset of the R_n, it is necessary and sufficient[17] that each group of $n + 2$ points contained in the set be congruent with $n + 2$ points of the R_n. Since $m > n + 3$, it follows that the m points may be imbedded isometrically in the R_n, and hence the determinant D vanishes.[18]

The order of D plays an essential rôle in Theorem 5.1. The hypotheses (i), (ii) are not sufficient to prove the vanishing of the determinant in case its order *equals* $n + 4$. Determinants of this type [i.e., of order $n + 4$ and satisfying

[15] This fact was called to my attention by Mr. M. Ville.

[16] Menger, loc. cit.

[17] Menger, *New foundation of euclidean geometry*, loc. cit., p. 727.

[18] This is evident, since D represents, to within a constant factor, the square of the volume of the simplex determined by the m points. Since this simplex is degenerate (the m points being in R_n, $n < m - 3$), its volume is zero.

hypotheses (i), (ii)] are of great interest in the characterization of the n-dimensional euclidean space, but little is known about them for $n > 1$. For $n = 1$, the characterization of pseudo-linear quadruples[19] enables us to prove at once the following

THEOREM 5.2. *If the determinant* $D = \begin{vmatrix} 0 & 1 \\ 1 & r_{ij} \end{vmatrix}$ $(i, j = 1, 2, 3, 4)$ *has each of its bordered fourth-order principal minors equal to zero,*[20] *then either D vanishes or* $r_{12} = r_{34}$, $r_{23} = r_{14}$, $r_{13} = r_{24}$, *one of the positive numbers* $\sqrt{r_{12}}$, $\sqrt{r_{23}}$, $\sqrt{r_{13}}$ *is the sum of the other two, and* $D = -32r_{12}r_{23}r_{13}$.

It is easily shown that if a determinant D of order $n + 4$ satisfies hypotheses (i), (ii) of Theorem 5.1 and does not vanish, its sign is the sign of $(-1)^n$, but neither the value of such a determinant, nor the relations existing between its elements (both explicitly exhibited for the case $n = 1$ by Theorem 5.2) has as yet been obtained for $n > 1$.

UNIVERSITY OF VIENNA AND INSTITUTE FOR ADVANCED STUDY.

[19] Menger, Mathematische Annalen, loc. cit., p. 126.

[20] This is hypothesis (ii) for $n = 1$; hypothesis (i) is automatically satisfied, since $r_{ij} > 0$ $(i \neq j)$.

TEMPERATURE DISTRIBUTION IN A SLAB OF TWO LAYERS

By R. V. Churchill

The list of solved problems in one-dimensional heat conduction in composite walls does not seem to include the cases in which the initial temperature distribution is arbitrary. The case of the semi-infinite composite solid with an arbitrary initial temperature distribution has been treated recently by Lowan.[1] The solution of the corresponding problem for the wall of finite thickness seems desirable for the sake of completeness. It is solved here by application of the Laplace transformation.

The problem under consideration is that of finding the one-dimensional distribution of temperature in a slab consisting of two layers of different materials whose outer parallel faces are held at fixed temperatures, when the initial temperature distribution in each layer is arbitrarily given. Let the thickness of the layers be a, b, and let $x = 0$ be taken as the surface of separation. Then the boundary conditions on the temperatures $T_1(x, t)$, $T_2(x, t)$ in the two layers may be written

$$\begin{aligned} &T_1(-a, t) = 0, \qquad \lim_{t\to 0} T_1(x, t) = f(x), \qquad -a < x < 0, \\ &T_2(b, t) = c, \qquad \lim_{t\to 0} T_2(x, t) = g(x), \qquad 0 < x < b, \end{aligned} \tag{1}$$

$$T_1(0, t) = T_2(0,t), \qquad K_1\frac{\partial}{\partial x} T_1(x, t) = K_2\frac{\partial}{\partial x} T_2(x, t), \qquad x = 0, \tag{2}$$

where K_1, K_2 are the thermal conductivities of the two layers.

It is easily seen that the temperatures T_1 and T_2 can be obtained by the composition of known temperature formulas and a simpler unknown formula. Let each of the three pairs of temperature functions u_1, u_2, v_1, v_2 and w_1, w_2 satisfy the conditions (2), and let

$$\begin{aligned} &u_1(-a, t) = 0, \qquad v_1(-a, t) = 0, \qquad w_1(-a, t) = 0, \\ &u_2(b, t) = c, \qquad v_2(b, t) = 0, \qquad w_2(b, t) = 0, \\ &\lim_{t\to 0} u_1(x, t) = 0, \quad \lim_{t\to 0} v_1(x, t) = f(x), \quad \lim_{t\to 0} w_1(x, t) = 0, \quad -a < x < 0, \\ &\lim_{t\to 0} u_2(x, t) = 0, \quad \lim_{t\to 0} v_2(x, t) = 0, \quad \lim_{t\to 0} w_2(x, t) = g(x), \quad 0 < x < b. \end{aligned}$$

Then the temperature functions

$$T_1 = u_1 + v_1 + w_1, \qquad T_2 = u_2 + v_2 + w_2$$

Received October 28, 1935.

[1] A. N. Lowan, *Heat conduction in a semi-infinite solid of two different materials*, this Journal, vol. 1 (1935), pp. 94–102.

satisfy the conditions (1) and (2). But the solution of the problem of finding the temperatures $u_1(x, t)$ and $u_2(x, t)$ is known,[2] and it is evident that the formulas for the functions w_1 and w_2 can be written at once from the solution of the problem of finding $v_1(x, t)$ and $v_2(x, t)$. Hence we shall consider the problem solved when we find the formulas for the temperatures $v_1(x, t)$ and $v_2(x, t)$.

The classical Fourier method of using an orthogonal set of functions leads to difficulties when applied to this problem. The method of the Laplace transformation which was introduced by Doetsch[3] enables us to transform our boundary-value problem into one in ordinary differential equations, whose solution is the Laplace transformation of the required solution. It has been shown by Carson[4] that this method is formally the same as that of using Heaviside's operators. A connection between the method of the Laplace transformation and Carslaw's[5] method of contour integrals will be seen at the end of this paper.

The Laplace transformation of a function $\phi(t)$ is defined by

$$L\{\phi(t)\} = \int_0^\infty e^{-st}\,\phi(t)dt = y(s). \tag{3}$$

The inverse of this operator is defined by the solution of the integral equation (3),

$$\phi(t) = L^{-1}\{y(s)\}.$$

In particular, for real positive values of λ and s, the integral

$$\int_0^\infty e^{-\lambda t}e^{-st}dt = \frac{1}{s+\lambda}$$

gives the useful result

$$L^{-1}\left\{\frac{1}{s+\lambda}\right\} = e^{-\lambda t}. \tag{4}$$

By integration by parts it is seen that the transformation (3) has the property

$$L\left\{\frac{d}{dt}\phi(t)\right\} = sL\{\phi(t)\} - \phi(0) = sy(s) - \phi(0). \tag{5}$$

[2] H. S. Carslaw, *Introduction to the Mathematical Theory of the Conduction of Heat in Solids*, 1921, p. 215. Carslaw's solution, which he obtains by using contour integrals, can be found easily by using the Laplace transformation.

[3] G. Doetsch, *Ueber das Problem der Wärmeleitung*, Jahresber. deutsch. Math. Ver., vol. 33 (1924), pp. 45–52. F. Bernstein and G. Doetsch, *Probleme aus der Theorie der Wärmeleitung* I, Math. Zeits., vol. 22 (1925), pp. 285–292, the first of a series of papers by these authors. Math. Zeits., vol. 22 (1925), pp. 293–306, vol. 25 (1926), pp. 608–626, vol. 26 (1927), pp. 89–98, vol. 28 (1928), pp. 567–578.

[4] J. R. Carson, *Electric Circuit Theory and the Operational Calculus*, 1926.

[5] H. S. Carslaw, footnote 2, Chapter 11.

The function ϕ may involve a parameter and the operator L is commutable with differentiation with respect to it.[6]

The conditions which determine our temperatures $v_1(x, t)$ and $v_2(x, t)$ are

$$\frac{\partial}{\partial t} v_1(x, t) = k_1 \frac{\partial^2}{\partial x^2} v_1(x, t) \qquad -a < x < 0, \tag{6}$$

$$v_1(-a, t) = 0, \tag{7}$$

$$\lim_{t \to 0} v_1(x, t) = f(x) \qquad -a < x < 0, \tag{8}$$

$$\frac{\partial}{\partial t} v_2(x, t) = k_2 \frac{\partial^2}{\partial x^2} v_2(x, t) \qquad 0 < x < b, \tag{9}$$

$$v_2(b, t) = 0, \tag{10}$$

$$\lim_{t \to 0} v_2(x, t) = 0 \qquad 0 < x < b, \tag{11}$$

$$v_1(0, t) = v_2(0, t), \tag{12}$$

$$K_1 \frac{\partial}{\partial x} v_1(x, t) = K_2 \frac{\partial}{\partial x} v_2(x, t) \qquad x = 0, \tag{13}$$

where k_1 and k_2 are the thermal diffusivities of the materials in the layers $-a < x < 0$ and $0 < x < b$, respectively.

Let the Laplace transformation (3) be applied to the members of equations (6) and (9). Then according to the property (5) and the conditions (8) and (11) the transformed functions

$$y_1(x, s) = L\{v_1(x, t)\}, \qquad y_2(x, s) = L\{v_2(x, t)\}$$

must satisfy the equations

$$k_1 \frac{\partial^2}{\partial x^2} y_1(x, s) = s y_1(x, s) - f(x) \qquad -a < x < 0, \tag{14}$$

$$k_2 \frac{\partial^2}{\partial x^2} y_2(x, s) = s y_2(x, s) \qquad 0 < x < b. \tag{15}$$

The applications of L to the conditions (7), (10), (12) and (13) give the following conditions, respectively:

$$y_1(-a, s) = 0, \tag{16}$$

$$y_2(b, s) = 0, \tag{17}$$

$$y_1(0, s) = y_2(0, s), \tag{18}$$

$$K_1 \frac{\partial}{\partial x} y_1(x, s) = K_2 \frac{\partial}{\partial x} y_2(x, s) \qquad x = 0. \tag{19}$$

[6] For other properties of the operator L and its inverse see reference (3), second paper, and G. Doetsch, *Die Integrodifferentialgleichungen vom Faltungstypus*, Math. Ann., vol. 89 (1923), pp. 192–207.

The equations (14) and (15) can be solved as ordinary linear differential equations involving the parameter s. Their general solutions can be written

$$y_1(x, s) = A_1 e^{x\sqrt{\frac{s}{k_1}}} + B_1 e^{-x\sqrt{\frac{s}{k_1}}} + \beta(x, s), \tag{20}$$

$$y_2(x, s) = A_2 e^{x\sqrt{\frac{s}{k_2}}} + B_2 e^{-x\sqrt{\frac{s}{k_2}}}, \tag{21}$$

where A_1, B_1, A_2, B_2 are arbitrary constants and

$$\beta(x, s) = \frac{1}{\sqrt{k_1 s}} \int_x^0 \sinh\left[\sqrt{\frac{s}{k_1}}\,(x - \xi)\right] \cdot f(\xi) d\xi. \tag{22}$$

When the constants are determined so that the conditions (16) to (19) are satisfied, the solutions (20), (21) can be written

$$\begin{aligned} y_1(x, s) = \beta(-a, s) \\ \frac{\sinh\left(x\sqrt{\frac{s}{k_1}}\right)\cosh\left(b\sqrt{\frac{s}{k_2}}\right) - \sigma\cosh\left(x\sqrt{\frac{s}{k_1}}\right)\sinh\left(b\sqrt{\frac{s}{k_2}}\right)}{\sinh\left(a\sqrt{\frac{s}{k_1}}\right)\cosh\left(b\sqrt{\frac{s}{k_2}}\right) + \sigma\cosh\left(a\sqrt{\frac{s}{k_1}}\right)\sinh\left(b\sqrt{\frac{s}{k_2}}\right)} + \beta(x, s), \end{aligned} \tag{23}$$

$$\begin{aligned} y_2(x, s) = -\sigma\beta(-a, s) \\ \frac{\sinh\left[\sqrt{\frac{s}{k_2}}\,(b - x)\right]}{\sinh\left(a\sqrt{\frac{s}{k_1}}\right)\cosh\left(b\sqrt{\frac{s}{k_2}}\right) + \sigma\cosh\left(a\sqrt{\frac{s}{k_1}}\right)\sinh\left(b\sqrt{\frac{s}{k_2}}\right)}, \end{aligned} \tag{24}$$

where

$$\sigma = \frac{K_1}{K_2}\sqrt{\frac{k_2}{k_1}}. \tag{25}$$

By changing to circular functions, equations (23) and (24) can be written

$$\begin{aligned} y_1(x, s) = \Big\{\phi(-a, s)\Big[\sin\left(x\sqrt{\frac{-s}{k_1}}\right)\cos\left(b\sqrt{\frac{-s}{k_2}}\right) \\ - \sigma\cos\left(x\sqrt{\frac{-s}{k_1}}\right)\sin\left(b\sqrt{\frac{-s}{k_2}}\right)\Big] + \phi(x, s)\, F\left(\sqrt{\frac{-s}{k_1}}\right)\Big\} \Big/ F\left(\sqrt{\frac{-s}{k_1}}\right), \end{aligned} \tag{26}$$

$$y_2(x, s) = -\frac{\sigma\,\phi(-a, s)\sin\left[\sqrt{\frac{-s}{k_2}}\,(b - x)\right]}{F\left(\sqrt{\frac{-s}{k_1}}\right)}, \tag{27}$$

where

$$\phi(x, s) = \frac{1}{\sqrt{-k_1 s}} \int_x^0 \sin\left[\sqrt{\frac{-s}{k_1}}(x - \xi)\right] f(\xi)\, d\xi = \beta(x, s) , \tag{28}$$

and the function F is defined by

$$F(\alpha) = \sin(a\alpha)\cos(b\mu\alpha) + \sigma\cos(a\alpha)\sin(b\mu\alpha) , \tag{29}$$

in terms of the thermal coefficient

$$\mu = \sqrt{\frac{k_1}{k_2}} . \tag{30}$$

Our problem is now resolved into one of finding the inverse Laplace transformations of these functions y_1 and y_2 given by equations (23) and (24) or (26) and (27), since

$$v_1(x, t) = L^{-1}\{y_1(x, s)\} , \qquad v_2(x, t) = L^{-1}\{y_2(x, s)\} .$$

We shall proceed by first obtaining a formal solution of this transformation problem by a method which is logical provided the series used have the necessary properties of convergence. The functions v_1 and v_2 so found will then be examined, with the aid of contour integrals, to show that they satisfy the required conditions.

In order to put the functions y_1 and y_2 into a form in which the operator L^{-1} may be applied we shall use the partial fractions theorem of Mittag-Leffler,[7] which shows how a meromorphic function $M(z)$ can be expressed as the sum of an integral function $G(z)$ (rational or transcendental) and a series which is determined by the poles and principal parts of $M(z)$. If $M(z)$ has only simple poles z_n ($n = 1, 2, 3, \cdots$) with residues ρ_n, the theorem shows that

$$M(z) = G(z) + \sum_{n=1}^{\infty}\left[\frac{\rho_n}{z - z_n} + P_n(z)\right],$$

where $P_n(z)$ are polynomials which may be necessary to satisfy conditions of convergence; the degree and coefficients of $P_n(z)$ depend upon ρ_n and z_n.

If the function $M(z)$ has the form $X(z)/Y(z)$, where $X(z)$ is integral and $Y(z)$ has an infinite number of zeros z_n, none of which are repeated [we exclude all z_n which are not poles of $X(z)/Y(z)$], the above expansion gives

$$\frac{X(z)}{Y(z)} = G(z) + \sum_{n=1}^{\infty} \frac{X(z_n)}{Y'(z_n)(z - z_n)} , \tag{31}$$

provided the conditions of convergence are satisfied with $P_n(z) \equiv 0$.

It has been shown[8] that the zeros α_n of the function $F(\alpha)$, defined by (29), are all real, infinite in number, and not repeated; it is evident that $-\alpha_n$ is a zero

[7] See for example E. J. Townsend, *Functions of a Complex Variable*, 1915, p. 303, or K. Knopp, *Funktionentheorie* II, (Sammlung Göschen, Nr. 703), p. 38.

[8] Carslaw, footnote 2, p. 215.

if α_n is a zero. It follows that the function $F(\sqrt{-s_n/k_1})$ in the denominators of (26) and (27) has an infinite number of distinct zeros s_n which are all real and negative except for $s = 0$, since

$$s_n = -k_1\,\alpha_n^2\,. \tag{32}$$

It is clear that $s = 0$ is not a pole of either $y_1(x, s)$ or $y_2(x, s)$; only poles of these functions are included in the set s_n.

Let the partial fractions expansion (31) be applied to the functions y_1 and y_2 in (26) and (27). Let $G_1(x, s)$ and $G_2(x, s)$ denote the unknown functions $G(z)$ which appear in these expansions; they are analytic throughout the finite s-plane. Also let $\bar{y}_1$ and $\bar{y}_2$ represent the partial fractions series part of the expansions, so that

$$y_1(x, s) = G_1(x, s) + \bar{y}_1(x, s)\,, \qquad y_2(x, s) = G_2(x, s) + \bar{y}_2(x, s)\,. \tag{33}$$

Then according to (31),

$$\begin{aligned}\bar{y}_1(x, s) = &-2\sqrt{k_1}\sum_{n=1}^{\infty}\sqrt{-s_n}\,\phi(-a, s_n)\\ &\frac{\sin\left(x\sqrt{\frac{-s_n}{k_1}}\right)\cos\left(b\sqrt{\frac{-s_n}{k_2}}\right)-\sigma\cos\left(x\sqrt{\frac{-s_n}{k_1}}\right)\sin\left(b\sqrt{\frac{-s_n}{k_2}}\right)}{F'\left(\sqrt{\frac{-s_n}{k_1}}\right)(s-s_n)}\\ = &-2\sqrt{k_1}\sum_{n=1}^{\infty}\frac{\sqrt{-s_n}\,\phi(-a, s_n)\sin\left[\sqrt{\frac{-s_n}{k_1}}\,(x+a)\right]\cos\left(b\sqrt{\frac{-s_n}{k_2}}\right)}{F'\left(\sqrt{\frac{-s_n}{k_1}}\right)(s-s_n)\cos\left(a\sqrt{\frac{-s_n}{k_1}}\right)}\,,\end{aligned} \tag{34}$$

since s_n are non-zero roots of

$$\begin{aligned}F\left(\sqrt{\frac{-s}{k_1}}\right) = \sin\left(a\sqrt{\frac{-s}{k_1}}\right)\cos\left(b\sqrt{\frac{-s}{k_2}}\right)\\ + \sigma\cos\left(a\sqrt{\frac{-s}{k_1}}\right)\sin\left(b\sqrt{\frac{-s}{k_2}}\right) = 0\,.\end{aligned} \tag{35}$$

Likewise,

$$\bar{y}_2(x, s) = 2\sigma\sqrt{k_1}\sum_{n=1}^{\infty}\frac{\sqrt{-s_n}\,\phi(-a, s_n)\sin\left[\sqrt{\frac{-s_n}{k_2}}\,(b-x)\right]}{F'\left(\sqrt{\frac{-s_n}{k_1}}\right)(s-s_n)}\,. \tag{36}$$

In order to determine the functions G_1 and G_2 we can make use of the conditions (14) to (19) which were used to determine y_1 and y_2. It is quite easily seen that $\bar{y}_1$ and $\bar{y}_2$ as given by equations (34) and (36) formally satisfy all the boundary conditions (16) to (19). Since y_1 and y_2 satisfy these conditions the corresponding conditions on G_1 and G_2 are determined, according to equations

(33). The formal substitution of the expression for $\bar{y}_1$ in (34) into the differential equation (14) leads to the equation

$$(37)\quad k_1 \frac{\partial^2}{\partial x^2}\bar{y}_1(x, s) - s\bar{y}_1(x, s)$$

$$= 2\sqrt{k_1}\sum_{n=1}^{\infty}\frac{\sqrt{-s_n}\,\phi(-a, s_n)\sin\left[\sqrt{\frac{-s_n}{k_1}}\,(x+a)\right]\cos\left(b\sqrt{\frac{-s_n}{k_2}}\right)}{F'\left(\sqrt{\frac{-s_n}{k_1}}\right)\cos\left(a\sqrt{\frac{-s_n}{k_2}}\right)}$$

$$= 2\sum_{n=1}^{\infty}\frac{\int_{-a}^{0}\sin\left[\alpha_n(a+\xi)\right]f(\xi)\,d\xi\,\sin\left[\alpha_n(a+x)\right]\cos\left(\alpha_n\,\mu b\right)}{F'(\alpha_n)\cos\left(\alpha_n\,a\right)},$$

where α_n are the positive zeros of $F(\alpha)$. But it is shown later on [see equations (44)–(48)], with the aid of contour integrals, that the series in (37) represents $f(x)$ in $(-a, 0)$. Hence $\bar{y}_1$ satisfies (14). Likewise,

$$(38)\quad \begin{aligned} &k_2\frac{\partial^2}{\partial x^2}\bar{y}_2(x, s) - s\,\bar{y}_2(x, s) \\ &\qquad = -2\sigma\sum_{n=1}^{\infty}\frac{\int_{-a}^{0}\sin\left[\alpha_n(a+\xi)\right]f(\xi)\,d\xi\,\sin\left[\alpha_n\,\mu(b-x)\right]}{F'(\alpha_n)},\end{aligned}$$

and this series represents a function which vanishes in $(0, b)$ [see equations (49)–(51)]. Since $\bar{y}_1$ and $\bar{y}_2$ satisfy all the conditions which determine y_1 and y_2, then $G_1 = G_2 = 0$, and

$$(39)\quad y_1(x, s) = \bar{y}_1(x, s)\,,\qquad y_2(x, s) = \bar{y}_2(x, s)\,.$$

This result was not entirely unexpected, since equations (34) and (36) indicate that $\bar{y}_1$ and $\bar{y}_2$ have the limit zero as s becomes infinite through positive values, and this property can be shown to apply to y_1 and y_2 by using the expressions (23) and (24).

When the inverse Laplace transformation is applied to the series in (34) and (36), the transformation (4) gives, in view of (39),

$$L^{-1}\{y_1(x, s)\}$$

$$= -2\sqrt{k_1}\sum_{n=1}^{\infty}\frac{\sqrt{-s_n}\,\phi(-a, s_n)\sin\left[\sqrt{\frac{-s_n}{k_1}}\,(x+a)\right]\cos\left(b\sqrt{\frac{-s_n}{k_2}}\right)}{F'\left(\sqrt{\frac{-s_n}{k_1}}\right)\cos\left(a\sqrt{\frac{-s_n}{k_1}}\right)}\,e^{s_n t},$$

$$L^{-1}y_2(x, s) = 2\sigma\sqrt{k_1}\sum_{n=1}^{\infty}\frac{\sqrt{-s_n}\,\phi(-a, s_n)\sin\left[\sqrt{\frac{-s_n}{k_2}}\,(b-x)\right]}{F'\left(\sqrt{\frac{-s_n}{k_1}}\right)}\,e^{s_n t}.$$

In terms of the positive roots $\alpha_n = \sqrt{-s_n/k_1}$ of the equation

$$(40) \qquad F(\alpha) = \sin(a\alpha)\cos(b\mu\alpha) + \sigma\cos(a\alpha)\sin(b\mu\alpha) = 0,$$

these equations can be written

$$(41) \quad v_1(x, t) = 2\sum_{n=1}^{\infty} \frac{\int_{-a}^{0} \sin[\alpha_n(a+\xi)]f(\xi)\,d\xi \,\sin[\alpha_n(a+x)]\cos(b\mu\alpha_n)}{F'(\alpha_n)\cos(a\alpha_n)} e^{-\alpha_n^2 k_1 t}$$

when $-a < x < 0$, and

$$(42) \quad v_2(x, t) = -2\sigma\sum_{n=1}^{\infty} \frac{\int_{-a}^{0} \sin[\alpha_n(a+\xi)]f(\xi)\,d\xi \,\sin[\alpha_n\mu(b-x)]}{F'(\alpha_n)} e^{-\alpha_n^2 k_1 t}$$

when $0 < x < b$.

It can be seen readily that our temperature formulas (41) and (42) formally satisfy all the conditions (6)–(13) on v_1 and v_2 except possibly for the initial temperature conditions (8) and (11). In the special case $\sigma = 1$, the series in (41) and (42) for $t = 0$ become Fourier sine series; the first represents a function which is equal to $f(x)$ in $(-a, 0)$ and vanishes in $(0, b\mu)$, and the second represents a function which equals $\sigma f(\mu x)$ in $(-a/\mu, 0)$ and vanishes in $(0, b)$. Hence in this case the initial conditions are satisfied. But except in such special cases the set of functions $\sin(\alpha_n x)$ is not orthogonal, so we shall make use of certain contour integrals which are suggested by the series in (41) and (42) to test our solution for the initial conditions.

Consider the contour integral

$$(43) \quad I_1(x, t) = \frac{1}{\pi i}\int_{(P)} \frac{\int_{-a}^{0} \sin[\alpha(a+\xi)]f(\xi)\,d\xi \,\sin[\alpha(a+x)]\cos(\alpha\mu b)}{F(\alpha)\cos(a\alpha)} e^{-\alpha^2 k_1 t}\,d\alpha,$$

where $-a \leqq x \leqq 0$, $t \geqq 0$, in the complex plane of $\alpha = \rho e^{i\theta}$ from right to left over the infinite path P consisting of the two half-lines $\theta = \pi/8$ and $\theta = 7\pi/8$, $(\rho > 0)$. (Because of the exponential factor it is essential that $\cos 2\theta > 0$ for large values of ρ on P.) After being multiplied by any fixed power of ρ the integrand of I_1 has the limit zero as ρ becomes infinite on the path P, provided that either x or t is not zero. It follows from the ordinary convergence test for infinite integrals that the integral (43) converges on each half of the path P for either $-a \leqq x \leqq 0$ and $t > 0$, or $-a \leqq x < 0$ and $t \geqq 0$. It may be noted that this is still true if the integrand is replaced by its derivative of any order with respect to either x or t.

The integral $I_1(x, t)$ is a continuous function of t at $t = 0$ if $-a \leqq x < 0$. This can be shown for each half of the path P, on which $\alpha = \rho e^{i\pi/8}$ and $\alpha = \rho e^{i7\pi/8}$, by writing the integrals in terms of ρ. The necessary properties of the real and imaginary parts of these integrals can be seen from those of the complex inte-

grands to show first that each integral is uniformly convergent in t for $0 \leqq t \leqq t'$, $x < 0$ and second, that the integrals are continuous. These results follow from two theorems on infinite integrals given by Carslaw.[9]

When we put $t = 0$, $-a \leqq x < 0$, in the integrand of (43) and integrate around the boundary of a circular sector above the radii $\theta = \pi/8$ and $\theta = 7\pi/8$ and under the circular arc $\rho = c$, the result is zero because all of the poles of the integrand lie on the axis of reals. But the integral over this circular arc approaches zero as the radius becomes infinite, and so the integral over the path P must vanish. Hence we can write

$$I_1(x, 0) = \lim_{t \to 0} I_1(x, t) = 0 \qquad -a \leqq x < 0. \tag{44}$$

To see the relation between $I_1(x, t)$ and $v_1(x, t)$, note first that the integrand in equation (43) is an odd function of α, so the integral over the path P is the same as the integral over the path made of the right halves of the lines $\theta = \pm\pi/8$, directed downward. Let an infinite sequence of circular arcs $\rho = \rho_k$ be used to join these rays and form circular sectors including a part of the positive real axis, just one definite ρ_k being selected between each pair of adjacent poles of the integrand, say. As ρ_k becomes infinite, it can be shown that the integral over the circular arc approaches zero. The sum of the integrals over the radii approaches $I_1(x, t)$. Since the set of poles α_n together with the positive roots of $\cos(\alpha a) = 0$ are the poles inclosed by this path, the theory of residues gives the result

$$I_1(x, t) = 2 \sum_{k=1}^{\infty} \frac{\int_{-a}^{0} \sin[\alpha_k(a + \xi)] f(\xi)\, d(\xi) \sin[\alpha_k(a + x)] \cos(\alpha_k \mu b) e^{-\alpha_k^2 k_1 t}}{F'(\alpha_k) \cos(\alpha_k a) - a F(\alpha_k) \sin(\alpha_k a)}, \tag{45}$$

where α_k are the positive roots of

$$F(\alpha) \cos(\alpha a) = 0.$$

Let λ_m denote the positive zeros of $\cos(\alpha a)$, so that

$$\lambda_m = (2m + 1)\pi/2a \qquad (m = 0, 1, 2, \cdots), \tag{46}$$

while α_n denotes the positive zeros of $F(\alpha)$. One term in each denominator of the series (45) is always zero, and when we separate the terms in which $\alpha_k = \lambda_m$ and simplify, equation (45) becomes

$$I_1(x, t) = 2 \sum_{n=1}^{\infty} \frac{\int_{-a}^{0} \sin[\alpha_n(a + \xi)] f(\xi) d\xi \sin[\alpha_n(a + x)] \cos(\alpha_n \mu b) e^{-\alpha_n^2 k_1 t}}{F'(\alpha_n) \cos(\alpha_n a)}$$
$$- \frac{2}{a} \sum_{m=0}^{\infty} \int_{-a}^{0} \cos(\lambda_m \xi) f(\xi) d\xi \cos(\lambda_m x) e^{-\lambda_m^2 k_1 t}.$$

[9] H. S. Carslaw, *Introduction to the Theory of Fourier Series and Integrals*, 1930, p. 196 and p. 198.

It follows from our expression (41) for $v_1(x, t)$ that

$$(47)\qquad v_1(x, t) = I_1(x, t) + \frac{2}{a}\sum_{m=0}^{\infty}\int_{-a}^{0}\cos(\lambda_m\xi)f(\xi)d\xi\,\cos(\lambda_m x)e^{-\lambda_m^2 k_1 t}.$$

According to equations (44) and (47) then,

$$\lim_{t\to 0} v_1(x, t) = \frac{2}{a}\sum_{m=0}^{\infty}\int_{-a}^{0}\cos(\lambda_m\xi)f(\xi)d\xi\,\cos(\lambda_m x).$$

The series on the right is a Fourier cosine series representing $f(x)$ in $(-a, 0)$, and hence our function $v_1(x, t)$ does satisfy the initial condition

$$(48)\qquad \lim_{t\to 0} v_1(x, t) = f(x) \qquad -a < x < 0.$$

By using the same paths of integration as before it can be shown in the same way that the contour integral

$$(49)\qquad I_2(x, t) = -\frac{\sigma}{\pi i}\int_{(P)}\frac{\displaystyle\int_{-a}^{0}\sin\,[\alpha(a+\xi)]\,f(\xi)d\xi\,\sin\,[\alpha\mu(b-x)]\,e^{-\alpha^2 k_1 t}d\alpha}{F(\alpha)}$$

converges and is represented by the series (42), so that

$$(50)\qquad v_2(x, t) = I_2(x, t) \qquad 0 < x < b.$$

This integral also approaches zero with t so that our initial condition

$$(51)\qquad \lim_{t\to 0} v_2(x, t) = 0 \qquad 0 < x < b,$$

is satisfied. Our functions $v_1(x, t)$ and $v_2(x, t)$ therefore satisfy all the initial and boundary conditions.

Since differentiation of the integrands of $I_1(x, t)$ and $I_2(x, t)$ once with respect to t or twice with respect to x introduces a factor α or α^2 as the essential change, it is readily shown that the integrals of these derivatives over the path P converge uniformly. Hence $I_1(x, t)$ and $I_2(x, t)$ can be differentiated inside the integral sign, and they satisfy the heat equations. It can be seen, then, that the expressions (47) and (50) for $v_1(x, t)$ and $v_2(x, t)$ in terms of contour integrals satisfy all the conditions of our problem. But it is not easy to see how we would arrive at this contour integral solution in this case without the use of some other method such as the one used here.

The solution of our problem then is given in series form by the equations (41) and (42), and in terms of contour integrals by (47) and (50). The latter form is especially useful for examining convergence properties.

UNIVERSITY OF MICHIGAN.

ASYMPTOTIC LINES THROUGH A PLANAR POINT OF A SURFACE AND LINES OF CURVATURE THROUGH AN UMBILIC

By Thomas L. Downs, Jr.

1. **Introduction.** If a regular point of a surface is not an umbilic, there pass through it two asymptotic lines (which may be imaginary) and two real lines of curvature.[1] If a point is an umbilic, this conclusion does not follow, for a circular umbilic is a singular point for the differential equation of the lines of curvature and a planar umbilic is a singular point for the differential equations of both families of curves. In a previous paper,[2] the author has studied the relations between two finite sets of directions at a planar point: the "true asymptotic directions", which are the possible tangent directions of the asymptotic lines through the point, and the "true principal directions", which are the possible tangent directions of the lines of curvature. In the present paper, we shall consider the asymptotic lines which are tangent to a given true asymptotic direction at a planar point and the lines of curvature which are tangent to a given true principal direction at a planar or circular point.

No previous results for the asymptotic lines appear to be known. The lines of curvature have been studied in the general case by Delloue;[3] the present paper amplifies his conclusions. In special cases, the lines of curvature have been treated by several writers, among whom may be mentioned Cayley,[4] Darboux,[5] Picard[6] and Wahlgren.[7]

The results for the asymptotic lines are not parallel to those for the lines of curvature. To an arbitrary true asymptotic direction at a planar point, there is tangent a unique asymptotic line in the most general case, two asymptotic lines in the next most general case. To an arbitrary true principal direction at a planar or circular point, there is tangent in general either a single line of curvature or an infinite number of lines of curvature, depending upon certain definite conditions.

Received December 24, 1935; part of a thesis presented at Harvard University.

[1] An umbilic is a regular point of a surface at which $e/E = f/F = g/G = 1/\rho$, where E, F, G and e, f, g are respectively the coefficients of the first and second fundamental forms of the surface. In the present paper an umbilic will be called a *circular point* if $1/\rho \neq 0$ and a *planar point* if $1/\rho = 0$.

[2] Downs, *Asymptotic and principal directions at a planar point of a surface*, this Journal, vol. 1 (1935), pp. 316–327.

[3] Comptes Rendus, vol. 187 (1928), p. 702.

[4] *On differential equations and umbilici*, Collected Math. Papers, vol. 5, p. 708.

[5] *Leçons sur la Théorie Générale des Surfaces*, vol. 4, Note VII.

[6] *Traité d'Analyse*, vol. 3, chap. IX, §14.

[7] Arkiv for Mat., Astr. och Fys., vol. 1 (1903), p. 43.

2. **The method of reduction.** The problem at hand is a special case of the general problem of analysis which seeks to determine the integral curves of a quadratic differential equation

$$A(x, y)dx^2 + 2B(x, y)dxdy + C(x, y)dy^2 = 0 , \tag{1}$$

which pass through a point at which all three coefficients vanish—a singular point. Let the origin be taken at the singular point of equation (1), and let us suppose that the coefficients A, B, C are analytic functions of (x, y) in some neighborhood of the origin. Then the differential equation takes the form

$$\begin{aligned}[A_n(x, y) + \cdots]\, dx^2 + 2\,[B_n(x, y) + \cdots]\, dxdy \\ + [C_n(x, y) + \cdots]\, dy^2 = 0 , \qquad n \geqq 1 ,\end{aligned} \tag{1a}$$

where A_n, B_n, C_n, homogeneous polynomials of degree n, not all identically zero, are the terms of lowest order in the Taylor expansions of the coefficients about the origin.

We shall consider only integral curves of (1a) which approach the origin so as to have a definite tangent there, that is, so that dy/dx approaches a definite limit, finite or infinite. Integral curves of this kind must be tangent at the origin to one of the $n + 2$ lines

$$A_n(x, y)\, x^2 + 2\, B_n(x, y)\, xy + C_n(x, y)\, y^2 = 0 . \tag{2}$$

Let $y/x = \lambda$ be a solution of the equation (2); we may suppose that λ is finite. Then if we set $y = (t + \lambda)x$ in (1a) and divide out x^n, we shall get a quadratic differential equation with a singular point at the origin in the (x, t)-plane. This equation is to be treated as we have just treated (1a).

A finite number of such substitutions suffices, at least in the cases of most frequent occurrence, to transform (1) into an equation

$$\alpha(\xi, \eta)d\xi^2 + 2\beta(\xi, \eta)d\xi d\eta + \gamma(\xi, \eta)d\eta^2 = 0 ,$$

whose discriminant

$$\Delta = \beta^2 - \alpha\gamma$$

does not vanish at $\xi = \eta = 0$, and which therefore factors into the two equations

$$\frac{d\eta}{d\xi} = \frac{a_1(\xi, \eta)}{b_1(\xi, \eta)} , \qquad \frac{d\eta}{d\xi} = \frac{a_2(\xi, \eta)}{b_2(\xi, \eta)} ,$$

in which a_i, b_i are analytic at $\xi = \eta = 0$. The theory of singular points of equations of this form is completely known. It is covered by a definitive paper of Bendixson's in the Acta.[8]

The method of reduction just described is a generalization of that used by Picard[9] in the special case when $n = 1$ and λ is a simple root of equation (2).

[8] Acta Math., vol. 24 (1900), pp. 1–88.

[9] *Traité d'Analyse*, vol. III, Chap. IX, p. 223 (1928 ed.).

Provided that λ is such a simple root, only one substitution is necessary no matter what the value of n, and the method of Picard applies essentially unchanged. A similar method for the treatment of the general case has also been suggested by Wahlgren.[10]

3. **The asymptotic lines through a planar point.** The asymptotic lines of an analytic surface

$$S: \qquad x_i = x_i(u, v) \qquad (i = 1, 2, 3)$$

are the integral curves of the differential equation

$$(3) \qquad e du^2 + 2 f du dv + g dv^2 = 0 ,$$

in which the left-hand side is the second fundamental form of S. The lines consist of two families of curves on S, one curve of each family passing through each regular non-planar point. A planar point is a singular point of (3), inasmuch as the coefficients e, f, g all vanish there.

If we choose the planar point P as the origin of rectangular coördinates (x, y, z) and the tangent plane to S at P as the (x, y)-plane, the surface in the neighborhood of P is represented by the equation

$$z = F(x, y) \equiv \varphi_n(x, y) + \varphi_{n+1}(x, y) + \cdots , \varphi_n(x, y) \not\equiv 0 , \quad n \geqq 3 ,$$

where $\varphi_k(x, y)$ is a homogeneous polynomial of degree k in x and y. If the coördinates (x, y) are regarded as the surface parameters, the differential equation (3) becomes

$$(3a) \qquad \left[\frac{\partial^2\varphi_n}{\partial x^2} + \frac{\partial^2\varphi_{n+1}}{\partial x^2} + \cdots\right] dx^2 + 2\left[\frac{\partial^2\varphi_n}{\partial x\partial y} + \frac{\partial^2\varphi_{n+1}}{\partial x\partial y} + \cdots\right] dx dy$$
$$+ \left[\frac{\partial^2\varphi_n}{\partial y^2} + \frac{\partial^2\varphi_{n+1}}{\partial y^2} + \cdots\right] dy^2 = 0 .$$

The integral curves of (3a) may be regarded as the orthogonal projections of the asymptotic lines upon the (x, y)-plane or as the asymptotic lines themselves referred to the surface coördinates $x = u$ and $y = v$. In either case their possible tangent directions at the origin—the true asymptotic directions at P — are defined by the equation

$$(4) \qquad \frac{\partial^2\varphi_n}{\partial x^2} x^2 + 2\frac{\partial^2\varphi_n}{\partial x\partial y} xy + \frac{\partial^2\varphi_n}{\partial y^2} y^2 \equiv n(n-1)\,\varphi_n(x, y) = 0 .$$

It will be convenient to use the slope-angle $\theta = \arctan y/x$ and to adopt the notation

$$\varphi_k(\theta) \equiv \varphi_k(\cos\theta, \sin\theta) , \qquad \varphi_k^{(p)}(\theta) \equiv \frac{d^p}{d\theta^p}\varphi_k(\cos\theta, \sin\theta) .$$

Then the following theorems are established by the method described.

[10] Bihang t. Kon. Svenska Vet.-Akad. Handlingar, vol. 28 (1902–1903), Afdelning 1, no. 4.

Let the angle α be a real root of order p of the equation

$$\varphi_n(\theta) = 0 . \tag{4a}$$

1. *If $\varphi_{n+1}(\alpha) \neq 0$, there is a unique asymptotic line tangent at the planar point to the direction $y:x = \tan \alpha$. It consists of two branches which join to form an analytic curve, smooth and without inflection at the origin if p is odd, but having a cusp of the first kind there if p is even. If $p = 1$, these conclusions hold even when $\varphi_{n+1}(\alpha) = 0$.*

2. *If $p > 1$ and if*

$$\varphi_{n+1}(\alpha) = 0 , \qquad \varphi'_{n+1}(\alpha) \neq 0 ,$$

then there are precisely two asymptotic lines tangent at the planar point to the direction $y:x = \tan \alpha$, unless

$$p = 2 , \qquad \Phi \geqq 1 ,$$

where

$$\Phi = \frac{2\, \varphi_{n+2}(\alpha) \cdot \varphi''_n(\alpha)}{[\varphi'_{n+1}(\alpha)]^2} .$$

In this case,

a) *if $\Phi = 1$, there is at least one asymptotic line in the given direction;*

b) *if $1 < \Phi \leqq \dfrac{(n^2 + 3n)^2}{(n^2 + 3n)^2 - 4}$, there is an infinite number;*

c) *if $\Phi > \dfrac{(n^2 + 3n)^2}{(n^2 + 3n)^2 - 4}$, there are none.*

To illustrate the method by which these results are established, we shall outline the proof of the first theorem. We suppose that $\varphi_{n+1}(\alpha) \neq 0$. If $p = 1$, the method of Picard applies and the one substitution $y = (t + \tan \alpha)x$ serves to establish the fact that there is a unique asymptotic line tangent to the direction of slope-angle α.

Let us then assume that $p > 1$; we may also suppose, without loss of generality, that $\alpha = 0$. Then

$$\varphi_n(1, t) = bt^p + \cdots \qquad b \neq 0 ,$$

and if we set $y = tx$ in the differential equation (3a) and divide through by bx^{n-2}, the equation becomes

$$\begin{aligned} &[t^p(A + \cdots) + x\,(a + \cdots) + x^2\,(\cdots)]\, dx^2 \\ &+ 2x\,[t^{p-1}(B + \cdots) + x\,(\cdots) + x^2\,(\cdots)]\, dx dt \\ &+ x^2\,[t^{p-2}(C + \cdots) + x\,(\cdots) + x^2\,(\cdots)]\, dt^2 = 0 , \end{aligned} \tag{5}$$

where $A = n(n - 1)$, $B = p(n - 1)$, $C = p(p - 1)$, and

$$a = \frac{p\,!\,n(n + 1)\, \varphi_{n+1}\,(1, 0)}{\dfrac{d^p}{dt^p}\,[\varphi_n\,(1, t)]_{t=0}} \neq 0 .$$

The possible tangent directions for solutions of (5) at the origin in the (x, t)-plane are given by the equation $ax^3 = 0$. Let us therefore set successively in (5) $x = x_1 t, x_1 = x_2 t, \cdots, x_{p-2} = x_{p-1} t$. We obtain a series of differential equations (I), (II), $\cdots$, (P − I) of which the k-th is

$$\begin{aligned} &t^2[t^{p-k}(A + \cdots) + ax_k + x_k t(\cdots)]\, dx_k^2 \\ \text{(K)} \qquad &+ 2tx_k\,[t^{p-k}(B_k + \cdots) + kax_k + x_k t(\cdots)]\, dx_k\, dt \\ &+ x_k^2\,[t^{p-k}(C_k + \cdots) + k^2 ax_k + x_k t(\cdots)]\, dt^2 = 0\,, \end{aligned}$$

where $B_k = kA + B, C_k = k^2A + 2kB + C$.

The tangent directions for solutions of (K) at the origin in the (x_k, t)-plane are defined by the equation

$$(\mathrm{E}_k) \qquad (k + 1)^2 ax_k^3\, t^2 = 0 \qquad (k < p - 1)\,.$$

Corresponding to the direction $t = 0$ of the equations $(\mathrm{E}_1), (\mathrm{E}_2), \cdots, (\mathrm{E}_{p-2})$ there will be found in each case only the solutions $t \equiv 0, x_k = t = 0$ of (K), each of which yields only the trivial solution $x = y = 0$ of the original equation (3a). On the other hand, the directions $x_k = 0$ defined by $(\mathrm{E}_1), (\mathrm{E}_2), \cdots, (\mathrm{E}_{p-2})$ have been disposed of by setting $x_k = x_{k+1} t$ in (K) and thus proceeding to the equation (K + I).

We consider now the differential equation (P − I); its solutions through the origin in the (x_{p-1}, t)-plane are tangent to the directions

$$(\mathrm{E}_{p-1}) \qquad x_{p-1}^2 t^2(p^2 ax_{p-1} + C_p t) = 0.$$

The root $t = 0$ of this equation is disposed of as we have done above. To discuss the root $x_{p-1} = 0$, we set in (P − I) $x_{p-1} = x_p t$ and obtain a differential equation (P) in x_p and t whose only solutions through the origin can be shown to be the axes $x_p \equiv 0$ and $t \equiv 0$; but these yield only the trivial solution $x = y = 0$ in the (x, y)-plane.

Thus we have left for consideration only the direction

$$p^2 ax_{p-1} + C_p t = 0$$

defined by (E_{p-1}). We therefore set in (P − I)

$$x_{p-1} = (u + \lambda)t, \qquad \lambda = -\frac{C_p}{p^2 a} \neq 0.$$

After division by t^2 there results the equation

$$\begin{aligned} &t^2[(A + a\lambda) + au + \cdots]\, du^2 \\ &+ 2t(u + \lambda)[(B_p + pa\lambda) + pau + \cdots]\, du dt \\ &+ (u + \lambda)^2[p^2 au + \cdots]\, dt^2 = 0. \end{aligned}$$

This equation factors into two equations of the first degree in du/dt:

$$(6a)\quad t\frac{du}{dt} = a_0 + \cdots, \qquad a_0 = \frac{-2(pn^2 + pn - p - 1)(pn^2 - 1)}{a(2pn - p - 1)} \neq 0;$$

$$(6b)\quad t\frac{du}{dt} = b_0 u + \cdots, \qquad b_0 = \frac{-p(pn^2 + pn - p - 1)}{2(pn^2 - 1)} < 0.$$

Equation (6a) has a regular point at $u = t = 0$ and so yields only the illusory solution $t \equiv 0$. To (6*b*) the criterion of Bendixson[11] applies: besides the illusory solution $t \equiv 0$ it has just one solution through the origin and tangent to the t-axis in the (u, t)-plane. This solution gives rise to a unique asymptotic line in the direction of the x-axis in the (x, y)-plane. It is shown by Picard[12] to be an analytic curve

$$u = ct^m + \cdots \qquad c \neq 0,\ m > 0,$$

in the (u, t)-plane. In the (x, y)-plane it is therefore represented by the equations

$$x = \lambda t^p + \cdots, \qquad y = \lambda t^{p+1} + \cdots,$$

and so has a cusp at the origin if p is even, but is smooth and without inflection if p is odd.

This completes the proof of the first theorem. The second is proved in a similar manner.

4. **The lines of curvature through an umbilic.** The lines of curvature on S are the integral curves of the differential equation

$$(7)\qquad \begin{vmatrix} edu + fdv & Edu + Fdv \\ fdu + gdv & Fdu + Gdv \end{vmatrix} = 0.$$

Let P be an umbilic on S, planar or circular, and choose coördinates as in §3 with P at the origin and the tangent plane to S at P as the (x, y)-plane. The surface in the neighborhood of P is then represented by the equation

$$S:\qquad z = \frac{1}{2\rho}(x^2 + y^2) + \varphi_3(x, y) + \cdots + \varphi_k(x, y) + \cdots,$$

where $1/\rho \neq 0$ or $1/\rho = 0$ according as P is a circular or planar point. If P is circular, the osculating sphere at P will be represented in the neighborhood of P by the equation

$$\Sigma:\qquad Z = \frac{1}{2\rho}(x^2 + y^2) + f_3(x, y) + \cdots + f_k(x, y) + \cdots,$$

[11] Loc. cit., p. 49.

[12] *Traité*, III (1928), p. 28 and p. 209.

where

$$f_{2k-1} \equiv 0, \qquad f_{2k} \equiv (a_{2k}/\rho^{2k-1})(x^2 + y^2)^k \qquad (a_{2k} \neq 0).$$

Let $Q(x, y, z)$ be a point on S near P. If P is circular, denote by $\bar{\varphi}_n(x, y)$ the principal part of the infinitesimal directed distance $z - Z$ from Q to Σ measured along the perpendicular from Q to the tangent plane at P:

$$\bar{\varphi}_n(x, y) \equiv \varphi_n(x, y) - f_n(x, y);$$

if P is planar, denote by $\bar{\varphi}_n(x, y)$ the principal part of the directed distance z from Q to the tangent plane at P. It is to be noted that if P is planar, $\bar{\varphi}_n \equiv \varphi_n$, where φ_n has the same meaning as in §3, and that in any case[13]

$$\bar{\varphi}_n(\theta) \equiv \varphi_n(\theta) + \text{constant}.$$

Under these conditions, the differential equation (7) of the lines of curvature takes the form[14]

$$\text{(7a)} \quad \begin{aligned} &\left[\frac{\partial^2\varphi_n}{\partial x\partial y} + \frac{\partial^2\varphi_{n+1}}{\partial x\partial y} + \cdots\right]dx^2 + \left[\left(\frac{\partial^2\varphi_n}{\partial y^2} - \frac{\partial^2\varphi_n}{\partial x^2}\right)\right. \\ &\left. + \left(\frac{\partial^2\varphi_{n+1}}{\partial y^2} - \frac{\partial^2\varphi_{n+1}}{\partial x^2}\right) + \cdots\right]dxdy - \left[\frac{\partial^2\varphi_n}{\partial x\partial y} + \frac{\partial^2\varphi_{n+1}}{\partial x\partial y} + \cdots\right]dy^2 = 0. \end{aligned}$$

When we set $x = r\cos\theta$, $y = r\sin\theta$, equation (7a) becomes

$$\text{(7b)} \quad \begin{aligned} &[\psi_n(\theta) + r\psi_{n+1}(\theta) + \cdots]dr^2 + r[\chi_n(\theta) + r\chi_{n+1}(\theta) + \cdots]drd\theta \\ &\qquad - r^2[\psi_n(\theta) + r\psi_{n+1}(\theta) + \cdots]d\theta^2 = 0 \qquad (n \geqq 3), \end{aligned}$$

where

$$\psi_k(\theta) \equiv (k-1)\varphi_k'(\theta) \equiv (k-1)\bar{\varphi}_k'(\theta), \qquad \chi_k(\theta) \equiv \varphi_k''(\theta) - k(k-2)\varphi_k(\theta).$$

We assume that ψ_n does not vanish identically, that is, that

$$\bar{\varphi}_n(x, y) \not\equiv a(x^2 + y^2)^m \qquad (2m = n).$$

The possible tangent directions at the origin for the integral curves of equation (7b)—the true principal directions at P—are then defined by the equation

$$\text{(8)} \qquad \psi_n(\theta) \equiv \psi_n(\cos\theta, \sin\theta) = 0.$$

[13] Downs, loc. cit., §7. There the directions defined by the equation $\bar{\varphi}_n = 0$ are called the "true osculatory directions at P".

[14] The omitted terms of the coefficients are of order at least n. This rule of formation holds only through terms of order $(n-1)$ if P is circular and through terms of order $(3n-5)$ if P is planar.

The results in the general case have been stated without proof by Delloue.[15] We shall here only sharpen the statement of his Case II by distinguishing two important subcases. Employing the methods already outlined, we arrive at the following theorem.

Suppose that γ *is a real root of order* p *of* $\psi_n(\theta)$ *and also a root of* $\chi_n(\theta)$, *but that* $\chi_n(\theta)$ *does not vanish identically.*

1. *If* $p = 1$ *and*

$$\text{(A)} \qquad [\psi_{n+1}\chi_n' - \psi_n'\chi_{n+1}]_{\theta=\gamma} \neq 0,$$

there is an infinite number of lines of curvature tangent at P *to the direction* $y:x = \tan\gamma$. *If* γ *is a multiple root of* $\chi_n(\theta)$, *the condition* (A) *becomes simply*

$$\text{(A}'\text{)} \qquad \chi_{n+1}(\gamma) \neq 0.$$

2. *If* $p \geqq 2$ *and* $\psi_{n+1}(\gamma) \neq 0$, *there is one and only one line of curvature tangent at* P *to the direction* $y:x = \tan\gamma$.

HARVARD UNIVERSITY.

[15] Loc. cit. His results are as follows in the notation of the present paper:

"I. Let OT be a real ray of the pencil $\psi_n(x, y) = 0$ which does not also belong to the pencil whose equation is $\chi_n(x, y) = 0$; γ, the angle it makes with Ox; $\psi_n^{(p)}(\theta)$ the first derivative of the function $\psi_n(\theta)$ which does not vanish for $\theta = \gamma$. If p is odd and if

$$\psi_n^{(p)}(\gamma)/\chi_n(\gamma) > 0,$$

there is one and only one line of curvature tangent to OT at O. In all other cases, there is an infinite number. There is always a real direction to which is tangent only a single line of curvature. To every real direction of the pencil $\psi_n(x, y) = 0$ there is tangent an analytic line of curvature (C), except in certain cases when $(n-1)\psi_n'(\gamma)/\chi_n(\gamma)$ is a negative integer. If between two consecutive lines (C) there are an infinite number of lines of curvature passing through the umbilic, the latter all have the same tangent there.

"II. If OT belongs to both the pencils $\psi_n(x, y) = 0$ and $\chi_n(x, y) = 0$, there is only a single line of curvature tangent to that ray at O in the cases of most frequent occurrence.

"III. The preceding conclusions suppose that $\bar{\varphi}_n(x, y)$ is not of the form $a(x^2 + y^2)^m$, $2m = n$. In that case an infinite number of lines of curvature with distinct tangents pass through the umbilic."

FORMAL PROPERTIES OF ORTHOGONAL POLYNOMIALS IN TWO VARIABLES

BY DUNHAM JACKSON

1. Construction and properties of symmetry of systems of orthogonal polynomials. The theory of orthogonal polynomials in two variables, as might be anticipated, presents numerous analogies with the corresponding theory in one variable, together with extensive and fundamental differences and complications, which add materially to the interest of the problem, and at the same time limit the scope of an elementary treatment of it.[1]

The "Schmidt process of orthogonalization" is applicable to functions of an arbitrary number of variables. If $\varphi_0(x, y)$, $\varphi_1(x, y)$, $\varphi_2(x, y)$, $\cdots$ form a set of functions integrable with their squares over a region R with no relation of linear dependence connecting any finite number of them (either identically or almost everywhere), it is possible to form a normalized orthogonal sequence $\Phi_0(x, y)$, $\Phi_1(x, y)$, $\Phi_2(x, y)$, $\cdots$ in which Φ_n is a linear combination of $\varphi_0, \varphi_1, \cdots, \varphi_n$. In particular, if R is finite and if $\rho(x, y)$ is a non-negative integrable function having a positive integral over R, application of the process to the linearly independent functions $\rho^{\frac{1}{2}}$, $\rho^{\frac{1}{2}}x$, $\rho^{\frac{1}{2}}y$, $\rho^{\frac{1}{2}}x^2$, $\rho^{\frac{1}{2}}xy$, $\rho^{\frac{1}{2}}y^2$, $\cdots$, taken in this order, gives a sequence of polynomials $q_{nm}(x, y)$, $n = 0, 1, 2, \cdots$; $m = 0, 1, \cdots, n$, such that

$$\iint_R \rho(x, y) q_{kl}(x, y) q_{nm}(x, y)\,dx\,dy = 0, \qquad |n - k| + |m - l| \neq 0,$$

$$\iint_R \rho(x, y)\,[q_{nm}(x, y)]^2 dx\,dy = 1.$$

The $n + 1$ polynomials $q_{n0}, q_{n1}, \cdots, q_{nn}$ are of the n-th degree in the two variables together, and with respect to ρ as weight function they are orthogonal

Received February 6, 1936; presented to the American Mathematical Society January 1, 1936.

[1] I am indebted to Professor Shohat for the following bibliographical indications: J. Shohat, *Théorie générale des polynomes orthogonaux de Tchebichef*, Mémorial des Sciences Mathématiques, No. 66, Paris, 1934, pp. 20–22, and references 25, 28, 29; F. Didon, *Étude de certaines fonctions analogues aux fonctions X_n de Legendre*, etc., Annales de l'École Normale Supérieure, vol. 5 (1868), pp. 229–310; F. Didon, *Développements sur certaines séries de polynomes*, ibid., vol. 7 (1870), pp. 247–268, and other articles by the same author in vols. 6 and 7 of the same Annales; P. Appell, *Sur une classe de polynomes à deux variables et le calcul approché des intégrales doubles*, Annales de la Faculté de Toulouse, vol. 4 (1890), pp. H 1–20; P. Appell, *Sur les fonctions hypergéométriques de plusieurs variables, les polynomes d'Hermite, et autres fonctions sphériques dans l'hyperespace*, Mémorial des Sciences Mathématiques, No. 3, Paris, 1925; P. Appell and J. Kampé de Fériet, *Fonctions hypergéométriques et hypersphériques—Polynomes d'Hermite*, Paris, 1926.

to each other and to every polynomial of lower degree; in these particular polynomials, though not generally in the case of the others presently to be introduced, the second subscript indicates the degree with respect to y. If polynomials $p_{n0}, p_{n1}, \cdots, p_{nn}$ are defined in terms of $q_{n0}, q_{n1}, \cdots, q_{nn}$ by a real orthogonal transformation of the form

$$p_{ni} = \sum_{j=0}^{n} c_{ij} q_{nj}, \qquad \sum_{j=0}^{n} c_{ij} c_{kj} = 0 \ (i \neq k), \qquad \sum_{j=0}^{n} c_{ij}^2 = 1$$

(whether with determinant 1 or with determinant -1), the polynomials p_{ni} are likewise of the n-th degree, orthogonal[2] to each other, normalized, and orthogonal to every polynomial of lower degree.

Any polynomial $P_n(x, y)$ of the n-th degree which is orthogonal to every polynomial of lower degree is necessarily a linear combination[3] of the $n + 1$ polynomials $q_{n0}, \cdots, q_{nn}$. For if the coefficient of y^n in $P_n(x, y)$ is c_n times the non-vanishing coefficient of y^n in $q_{nn}(x, y)$, the polynomial $P_n(x, y) - c_n q_{nn}(x, y)$ has no term in y^n, and is still orthogonal to every polynomial of lower degree; if the coefficient of xy^{n-1} in $P_n - c_n q_{nn}$ is c_{n-1} times the coefficient of xy^{n-1} in $q_{n,n-1}$, terms in y^n and xy^{n-1} are both absent from the remainder

$$P_n - c_n q_{nn} - c_{n-1} q_{n,n-1};$$

and by continuation of the indicated process there is obtained ultimately a polynomial

$$P_n - c_n q_{nn} - c_{n-1} q_{n,n-1} - \cdots - c_{n0} q_{n0}$$

which contains no term of the n-th degree, but is orthogonal to every polynomial of degree lower than the n-th, and so in particular must be orthogonal to itself, and hence identically zero. It is equally true that P_n can be linearly expressed in terms of any set $p_{n0}, \cdots, p_{nn}$ defined as in the preceding paragraph, since the q's can be expressed in terms of the p's.

If $\pi_{n0}(x, y), \cdots, \pi_{nn}(x, y)$ is any set of $n + 1$ normalized orthogonal polynomials of the n-th degree orthogonal to every polynomial of lower degree, the π's are expressible in terms of the p's by an orthogonal transformation. For each π_{ni}, as just noted, can be written in the form

$$\pi_{ni} = \sum_{j=0}^{n} \gamma_{ij} p_{nj},$$

[2] This term as applied to pairs of polynomials will be understood in each case to mean orthogonal with respect to the weight function under consideration; and a corresponding interpretation is to be attached to the word *normalized*.

[3] This is of course not true in general of an arbitrary polynomial of the n-th degree, which involves $(n + 1)(n + 2)/2$ coefficients.

and by the assumption that the π's are normalized and orthogonal

$$0 = \iint_R \rho(x, y)\pi_{ni}(x, y)\pi_{nk}(x, y)dx\,dy = \sum_{j=0}^{n} \gamma_{ij}\gamma_{kj} \qquad (i \neq k),$$

$$1 = \iint_R \rho(x, y)\,[\pi_{ni}(x, y)]^2 dx\,dy = \sum_{j=0}^{n} \gamma_{i\,j}^2.$$

This leads to a relation between properties of symmetry of the weight function $\rho(x, y)$ and corresponding properties of the orthogonal polynomials. Let it be supposed that there is a transformation

$$x' = Ax + By, \qquad y' = Cx + Dy, \tag{1}$$

(necessarily[4] of determinant ± 1), which carries the region R into itself, and under which furthermore the function ρ is invariant: $\rho(x', y') \equiv \rho(x, y)$. Let $p(x', y')$ be a polynomial of the n-th degree which is orthogonal to every polynomial of lower degree, and let $p(x', y') \equiv \pi(x, y)$. Let $\sigma(x, y) \equiv s(x', y')$ be an arbitrary polynomial of degree $n - 1$ at most; the degree is naturally the same with respect to either pair of variables. Then

$$\iint \rho(x, y)\pi(x, y)\sigma(x, y)dx\,dy = \iint_R \rho(x', y')p(x', y')s(x', y')dx'dy' = 0. \tag{2}$$

In view of the arbitrariness of $\sigma(x, y)$ this means that $\pi(x, y)$ is a linear combination of $p_{n0}(x, y), \cdots, p_{nn}(x, y)$. If $p(x, y)$ is normalized, the determinant of the transformation being ± 1,

$$\iint_R \rho(x, y)\,[\pi(x, y)]^2 dx\,dy = \iint_R \rho(x', y')\,[p(x', y')]^2 dx'\,dy' = 1,$$

and $\pi(x, y)$ is normalized also. In (2) the polynomials p, s can be replaced by any two polynomials which are orthogonal to each other, in particular by any two of the polynomials $p_{n0}, \cdots, p_{nn}$. If

$$\pi_{n0}(x, y) \equiv p_{n0}(x', y'), \cdots, \pi_{nn}(x, y) \equiv p_{nn}(x', y'),$$

the π's form a system of normalized orthogonal polynomials of the n-th degree, orthogonal to every polynomial of lower degree, to which the preceding paragraph is applicable. *A transformation* (1) *under which R and ρ are invariant defines an orthogonal transformation of the set of polynomials $p_{n0}, \cdots, p_{nn}$, for*

[4] If J is the determinant of the transformation, and E the area of the region,

$$E = \iint_R dx'dy' = |J| \iint_R dx\,dy = |J|\,E.$$

The transformation on x, y is, however, not necessarily unitary; e.g., the transformation $x' = 2y$, $y' = \frac{1}{2}x$ carries the rectangle $-2 \leqq x \leqq 2$, $-1 \leqq y \leqq 1$ into the rectangle $-2 \leqq x' \leqq 2$, $-1 \leqq y' \leqq 1$, and leaves the positive function $1 + x^2 + 4y^2$ invariant.

each value of n. To a group of transformations (1) corresponds for each n a simply or multiply isomorphic group of transformations[5] of the p's.

For example, if R is the square $-1 \leqq x \leqq 1$, $-1 \leqq y \leqq 1$, if $\rho(x, y)$ has the form $\rho_1(x)\rho_2(y)$, and if $p_0(x)$, $p_1(x)$, $\cdots$ and $q_0(y)$, $q_1(y)$, $\cdots$ are the systems of normalized orthogonal polynomials in one variable corresponding to the weight functions ρ_1 and ρ_2, respectively, the polynomials $p_{n0}, \cdots, p_{nn}$ can be taken as

$$p_n(x)q_0(y),\ p_{n-1}(x)q_1(y),\ \cdots,\ p_0(x)q_n(y).$$

If ρ_1 and ρ_2 are even functions, the orthogonal polynomials in one variable are even or odd according as the degree is even or odd, and the transformation $x' = -x$, $y' = -y$ carries over the polynomial $p_{ni}(x, y) \equiv p_{n-i}(x)q_i(y)$ into

$$\pi_{ni}(x, y) = p_{n-i}(-x)q_i(-y) = (-1)^n p_{n-i}(x)q_i(y) = (-1)^n p_{ni}(x, y);$$

for n even the transformations of the group

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

on x and y both correspond to the identical transformation on $p_{n0}, \cdots, p_{nn}$, while for n odd they correspond respectively to the identical transformation and to its negative.[6] The group

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

on x and y gives rise to two or four different transformations on the p's according as n is even or odd. The group relationships which arise in various cases would obviously constitute an extensive study in themselves.

As another example, suppose that ρ is a symmetric function of x and y, $\rho(y, x) \equiv \rho(x, y)$, the region R being one which is carried over into itself by the transformation $x' = y$, $y' = x$. Let $q(x, y)$ be a polynomial of the n-th degree which is orthogonal to every polynomial of lower degree, and normalized,

$$\int\int_R \rho(x, y)\,[q(x, y)]^2 dx\, dy = 1.$$

By the general discussion above $q(y, x)$ is likewise normalized and orthogonal to every polynomial of lower degree. The polynomials $q(x, y) + q(y, x)$ and

[5] These facts are of course suggested by the applications of group theory in quantum mechanics. The writer is not aware that the present formulation is a familiar one.

[6] The same thing is readily seen to be true in the case of any weight function such that $\rho(-x, -y) \equiv \rho(x, y)$, with any choice of the polynomials p_{ni}. For the terms of the n-th degree in $p_{ni}(-x, -y)$ are the corresponding terms of $p_{ni}(x, y)$ multiplied by $(-1)^n$, and if $p_{ni}(-x, -y)$ as well as $p_{ni}(x, y)$ is orthogonal to every polynomial of lower degree, it follows that $p_{ni}(-x, -y) - (-1)^n p_{ni}(x, y)$, having no terms of the n-th degree, must be orthogonal to itself.

$q(x, y) - q(y, x)$, still orthogonal to every polynomial of lower degree, are also orthogonal to each other, since

$$\iint_R \rho(x, y)\,[q(x, y) + q(y, x)]\,[q(x, y) - q(y, x)]dx\,dy$$
$$= \iint_R \rho(x, y)\,[q(x, y)]^2dx\,dy - \iint_R \rho(x, y)\,[q(y, x)]^2dx\,dy = 1 - 1 = 0.$$

Unless one of them is identically zero, they can be normalized by means of the appropriate constant factors, that is to say, the polynomial $q(x, y)$ either is itself symmetric or skew-symmetric, or gives rise to a pair of polynomials of similar character of which one is symmetric and the other skew-symmetric.

If $r(x, y)$ is another polynomial of the n-th degree which is orthogonal to every polynomial of lower degree, and not linearly dependent on $q(x, y)$ and $q(y, x)$, it is possible to form a linear combination

$$q_1(x, y) = r(x, y) - cq(x, y) - c'q(y, x)$$

which is not identically zero and is orthogonal to $q(x, y)$ and to $q(y, x)$. Then $q_1(y, x)$ also is orthogonal to $q(x, y)$ and to $q(y, x)$, in consequence of the symmetry of ρ. This polynomial $q_1(x, y)$, like $q(x, y)$, either is symmetric or skew-symmetric or gives rise to a pair of orthogonal polynomials, one symmetric and the other skew-symmetric. If there is a polynomial $r_1(x, y)$ of the n-th degree orthogonal to every polynomial of lower degree and not linearly dependent on $q(x, y)$, $q(y, x)$, $q_1(x, y)$ and $q_1(y, x)$, the process can be continued. It leads ultimately to a set of $n + 1$ polynomials of the n-th degree, all symmetric or skew-symmetric, normalized, and orthogonal to each other as well as to every polynomial of lower degree.

If the construction is based on the particular set of polynomials $q_{n0}, \cdots, q_{nn}$ defined at the beginning of the paper, it can be said with definiteness that $q_{n0}(x, y)$ is neither symmetric nor skew-symmetric (for $n > 0$), since it contains a term in x^n and no term in y^n; the result of subtracting from $q_{n1}(x, y)$ a linear combination of $q_{n0}(x, y)$ and $q_{n0}(y, x)$ is neither symmetric nor skew-symmetric (for $n > 2$), since it contains a term in $x^{n-1}y$ and no term in xy^{n-1}, and so on. The resulting set consists of $\frac{1}{2}(n + 1)$ pairs, or of $\frac{1}{2}n$ pairs and a single polynomial, according as n is odd or even. In the latter case, the single polynomial obtained after construction of the $\frac{1}{2}n$ pairs must be symmetric, for it is certainly not skew-symmetric, since it actually contains a term in $x^{n/2}y^{n/2}$. *When $\rho(x, y)$ is symmetric, the polynomials $p_{n0}, \cdots, p_{nn}$ of the n-th degree in the orthogonal system can be chosen so that the matrix of the transformation by which the polynomials $p_{ni}(y, x)$ are expressed in terms of the polynomials $p_{ni}(x, y)$ is in diagonal form, with $\frac{1}{2}(n + 1)$ or $\frac{1}{2}n + 1$ of the diagonal elements equal to $+1$, according as n is odd or even, and the rest equal to -1.* The transformations of the p's corresponding to the group

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

on x and y are in this case distinct for every $n \geqq 1$.

Parts of the above reasoning are applicable under more general circumstances, or at any rate under different circumstances. Let a sequence of functions φ_k be given as at the beginning of the paper, and as they occur in order let them be grouped into sets in any way, with μ_1 functions in the first set, μ_2 in the second set, and so on. A linear combination of φ's involving one or more functions from the n-th set, with or without functions from earlier sets, but not containing any from sets beyond the n-th, will be called a *sum of the n-th grade*. Let μ_n for a particular value of n be represented for simplicity by the symbol ν. The Schmidt process, applied to the φ's in order, gives ν sums of the n-th grade, which may be denoted by $\Phi_{n1}, \Phi_{n2}, \cdots, \Phi_{n\nu}$, normalized and orthogonal to each other and orthogonal to every sum of lower grade. Any set of ν functions $\Psi_{n1}, \Psi_{n2}, \cdots, \Psi_{n\nu}$ expressed in terms of $\Phi_{n1}, \cdots, \Phi_{n\nu}$ by an orthogonal transformation of the form previously considered will be a set of normalized orthogonal sums of the n-th grade, orthogonal to every sum of lower grade. Any sum of the n-th grade which is orthogonal to every sum of lower grade is linearly expressible in terms of $\Phi_{n1}, \cdots, \Phi_{n\nu}$, for subtraction of suitable multiples of $\Phi_{n\nu}, \Phi_{n,\nu-1}, \cdots$ in succession leaves a remainder which contains no term of the n-th grade, and so must be orthogonal to itself and hence identically zero. Any two sets of normalized orthogonal sums of the n-th grade, orthogonal to every sum of lower grade, with ν sums in each set, must be expressible in terms of each other by orthogonal transformation.

These considerations apply to orthogonal polynomials classified otherwise than by the degree of the polynomial in the two variables jointly. If $\rho(x, y)$ is a given weight function as before, and if a polynomial is said to be of the n-th grade when the exponent of the highest power of either variable occurring in it is n, sets of orthogonal polynomials of grade $0, 1, 2, \cdots$ are obtained by applying the Schmidt process to the products of $\rho^{\frac{1}{2}}$ by the monomials $1, x, xy, y, x^2, x^2y, x^2y^2, xy^2, y^2, \cdots$ successively; the value of μ_n in this case is $2n + 1$. The scope of the earlier developments with regard to invariance of $\rho(x, y)$ under linear transformation of x and y is limited by the fact that such a transformation does not in general leave the grade of a polynomial unaltered; for example xy, which is of the first grade, is in general carried over into a polynomial of the second grade. Transformations of the form $x' = \pm x$, $y' = \pm y$, or of the form $x' = \pm y$, $y' = \pm x$, do however leave the grade of a polynomial unchanged, and if R and ρ are invariant under one of these transformations, it defines orthogonal transformations of the various sets of polynomials of specified grade in the orthogonal system. If $\rho(y, x) \equiv \rho(x, y)$, the $2n + 1$ polynomials of the n-th grade can be chosen for each value of n so that $n + 1$ of them are symmetric and the remaining n skew-symmetric.

With a corresponding definition of the grade of a trigonometric sum in two

variables, the grade of such a sum is unaltered by the transformations $x' = \pm x$, $y' = \pm y$ and $x' = \pm y$, $y' = \pm x$, as well as the order, considered to be the sum of the orders with respect to the two variables separately, and a theory of orthogonal transformation can be worked out for the corresponding sets of orthogonal sums.

To return to the case first considered, that of a system of orthogonal polynomials classified according to degree in the two variables together, if $q_{n0}, \cdots, q_{nn}$ are subjected to a complex unitary transformation, the resulting polynomials $p_{n0}, \cdots, p_{nn}$ (in which x and y are still to be thought of as real variables ranging over the region R) will be orthogonal to each other in the Hermitian sense, normalized, and orthogonal to every polynomial of lower degree (with real or complex coefficients) in the Hermitian sense as well as otherwise, since in the relationship with polynomials of lower degree the real and pure imaginary parts are orthogonal separately. So the real orthogonal transformation of the q's induced by a linear transformation of the form (1) which leaves R and ρ invariant can be reduced to normal form by suitable choice of the p's according to the general theory of unitary matrices, and even if the p's thus introduced are complex, they still have a definite significance as orthogonal polynomials.[7]

2. **Recursion formula and Christoffel-Darboux identity.** The processes which lead to the recursion formula and the Christoffel-Darboux identity for orthogonal polynomials in one variable are applicable also in the case of two variables, though the results are naturally less simple, and their utility for the theory of convergence of the corresponding developments in series is not so readily apparent.

Let a set of normalized orthogonal polynomials corresponding to a weight function $\rho(x, y)$ in a region R be denoted as before by $p_{ni}(x, y)$, $n = 0, 1, 2, \cdots$; $i = 0, 1, \cdots, n$, each polynomial being of the degree indicated by its first subscript. For specified n and i the product $xp_{ni}(x, y)$, being a polynomial of degree $n + 1$, can be expressed in the form

$$xp_{ni}(x, y) = \sum_{m=0}^{n+1} \sum_{j=0}^{m} c_{mj}p_{mj}(x, y),$$

with

$$c_{mj} = \int\int_R \rho(x, y)xp_{ni}(x, y)p_{mj}(x, y)dx\,dy.$$

This coefficient of course depends on n and i as well as m and j, but a corresponding elaboration of the symbolism is unnecessary. As $xp_{mj}(x, y)$ is a polynomial of degree $m + 1$, and as $p_{ni}(x, y)$ is orthogonal to every polynomial of degree lower than the n-th, $c_{mj} = 0$ if $m < n - 1$. So

$$xp_{ni} = \sum_{j=0}^{n+1} c_{n+1,j}p_{n+1,j} + \sum_{j=0}^{n} c_{nj}p_{nj} + \sum_{j=0}^{n-1} c_{n-1,j}p_{n-1,j}.$$

[7] The linear transformations discussed in this section have been further studied by Mr. Andrew Sobczyk in a master's thesis at the University of Minnesota.

The three sums on the right represent polynomials of degrees $n + 1$, n, and $n - 1$ respectively, each orthogonal to every polynomial of degree lower than its own. If these are normalized, and if the normalized polynomials are represented by $U_{n+1,i}(x, y)$, $V_{ni}(x, y)$, and $W_{n-1,i}(x, y)$, the identity takes the form

$$xp_{ni}(x, y) = \alpha_{ni}U_{n+1,i}(x, y) + \beta_{ni}V_{ni}(x, y) + \gamma_{ni}W_{n-1,i}(x, y).$$

(If one of the sums is identically zero, it can be regarded as the product of an arbitrary normalized polynomial of appropriate degree by a vanishing coefficient.) The normalized polynomials $U_{n+1,i}$, V_{ni}, $W_{n-1,i}$, of degrees $n + 1$, n, and $n - 1$, are each orthogonal to every polynomial of lower degree, and in particular, for fixed n and i, are orthogonal to each other, but (as far as appears from the present reasoning) it is not to be supposed that

$$U_{n+1,0},\ U_{n+1,1},\ \cdots,\ U_{n+1,n}$$

are in general orthogonal to each other, or that $V_{n0}, \cdots, V_{nn}$ are orthogonal among themselves, or $W_{n-1,0}, \cdots, W_{n-1,n}$, or that V_{ni} is the same as p_{ni}.

In consequence of the specified properties of the polynomials U, V, W

$$\iint_R \rho(x, y)[xp_{ni}(x, y)]^2 dx\, dy = \alpha_{ni}^2 + \beta_{ni}^2 + \gamma_{ni}^2.$$

On the other hand, if G is the greatest value of $|x|$ in R, this integral can not exceed G^2, since p_{ni} is normalized. So

$$\alpha_{ni}^2 + \beta_{ni}^2 + \gamma_{ni}^2 \leqq G^2,$$

where G depends only on R, and in particular is independent of n and i.

Similar reasoning is applicable to the product $yp_{ni}(x, y)$, or to

$$(Ax + By)p_{ni}(x, y),$$

if A and B are any constants.

The formulas thus obtained may be regarded collectively as corresponding to the recursion formula connecting successive members of a set of orthogonal polynomials in one variable.

An "arbitrary" function $f(x, y)$ can be formally expanded in a series of the form

$$\sum_{k=0}^{\infty} \sum_{i=0}^{k} c_{ki}\, p_{ki}(x, y),$$

with

$$c_{ki} = \iint_R \rho(u, v) f(u, v)\, p_{ki}(u, v) du\, dv.$$

If $S_n(x, y)$ denotes the partial sum of this series through terms of the n-th degree and if

$$K_n(x, y, u, v) \equiv \sum_{k=0}^{n} \sum_{i=0}^{k} p_{ki}(x, y) p_{ki}(u, v),$$

then

$$S_n(x, y) = \iint_R \rho(u, v) f(u, v) K_n(x, y, u, v) du\, dv.$$

If the polynomials $p_{k0}, \cdots, p_{kk}$ are replaced by an alternative set by means of an orthogonal transformation, the sum

$$\sum_{i=0}^{k} p_{ki}(x, y) p_{ki}(u, v)$$

is invariant under this transformation, and consequently the whole expression $K_n(x, y, u, v)$, in conformity with the fact that $S_n(x, y)$ is definable independently of any particular orthogonal system as that polynomial of the n-th degree (at most) for which

$$\iint_R \rho(x, y)[f(x, y) - S_n(x, y)]^2 dx\, dy$$

is a minimum.

The series expansion of a polynomial amounts to nothing more than a rearrangement of the polynomial itself, and any polynomial $P_n(x, y)$ of the n-th or lower degree is reproduced identically by the formula

$$P_n(x, y) = \iint_R \rho(u, v) P_n(u, v) K_n(x, y, u, v) du\, dv.$$

Considered as a function of u and v, the product $(u - x)K_n(x, y, u, v)$ is a polynomial of degree $n + 1$. As such it can be expressed in the form

$$\sum_{k=0}^{n+1} \sum_{i=0}^{k} c_{ki} p_{ki}(u, v),$$

in which the c's are functions of x and y given by

$$\begin{aligned} c_{ki} &= \iint_R \rho(u, v)(u - x) K_n(x, y, u, v) p_{ki}(u, v) du\, dv \\ &= \iint_R \rho u K_n p_{ki} du\, dv - x \iint_R \rho K_n p_{ki} du\, dv. \end{aligned}$$

For $k < n$, the function $up_{ki}(u, v)$ is a polynomial of the n-th or lower degree, and hence, by the preceding paragraph,

$$\iint_R \rho(u, v) u p_{ki}(u, v) K_n(x, y, u, v) du\, dv \equiv x p_{ki}(x, y),$$

while the integral with the factor u omitted is identically equal to $p_{ki}(x, y)$. So[8] $c_{ki} = 0$ for $k < n$.

[8] For a corresponding argument in one variable, ascribed to J. Geronimus, see J. Shohat, *On Stieltjes continued fractions*, American Journal of Mathematics, vol. 54 (1932), pp. 79–84; p. 81.

Any polynomial of degree $n + 1$ in u and v and of the same degree in x and y can be written in the form

$$\sum_{k=0}^{n+1} \sum_{i=0}^{k} \sum_{l=0}^{n+1} \sum_{j=0}^{l} c_{kilj}\, p_{lj}(x, y) p_{ki}(u, v).$$

If $(u - x)K_n(x, y, u, v)$ is so expressed, the fact that for $k < n$ the coefficient of $p_{ki}(u, v)$ is identically zero as a function of x and y means that all the coefficients c_{kilj} in which $k < n$ vanish.[9] By the skew-symmetry of the function for interchange of the pair of variables (u, v) with the pair (x, y) it appears that $c_{kilj} = 0$ also whenever $l < n$. When parts of the calculation are made more explicit, as will be done presently, everything being expressed in terms of p's, all terms which are of lower degree than the n-th in either pair of variables must cancel out in the final result, and such terms need not be traced in detail through the intermediate stages of the work.

It is sufficient accordingly to begin by observing that

$$(u - x)K_n(x, y, u, v) = (u - x) \sum_{i=0}^{n} p_{ni}(x, y) p_{ni}(u, v) + \text{terms of lower degree.}$$

By insertion of the integral expressions for the coefficients in the recursion formula

$$up_{ni}(u, v) = \sum_{j=0}^{n+1} \iint_R \rho(r, s) r p_{ni}(r, s) p_{n+1,j}(r, s) p_{n+1,j}(u, v)\,dr\,ds$$

$$+ \sum_{j=0}^{n} \iint_R \rho(r,s) r p_{ni}(r,s) p_{nj}(r,s) p_{nj}(u,v)\,dr\,ds + \text{terms of lower degree.}$$

Similarly, with an interchange of the subscripts i and j, which is arbitrary for the moment but does not affect the validity of the formula,

$$xp_{nj}(x, y) = \sum_{i=0}^{n+1} \iint_R \rho(r, s) r p_{nj}(r, s) p_{n+1,i}(r, s) p_{n+1,i}(x, y)\,dr\,ds$$

$$+ \sum_{i=0}^{n} \iint_R \rho(r, s) r p_{nj}(r, s) p_{ni}(r, s) p_{ni}(x, y)\,dr\,ds + \text{terms of lower degree.}$$

On multiplication of the identity for $up_{ni}(u, v)$ by $p_{ni}(x, y)$ and summation with respect to i, the terms of the n-th degree with respect to each pair of variables in the expansion of $u \sum_i p_{ni}(x, y) p_{ni}(u, v)$ are seen to be

$$\sum_{i=0}^{n} \sum_{j=0}^{n} \iint_R \rho(r, s) r p_{ni}(r, s) p_{nj}(r, s) p_{nj}(u, v) p_{ni}(x, y)\,dr\,ds,$$

[9] Here and elsewhere it is to be noted that linear independence of the p's, obvious from the manner of their construction, is also immediately deducible, without reference to the details of that process, from their property of orthogonality.

while the terms of like degree in the corresponding expression for

$$x \sum_j p_{nj}(x, y)p_{nj}(u, v)$$

are the same, so that these terms cancel out of the representation of

$$(u - x)K_n(x, y, u, v),$$

together with all terms of lower degree, and only the terms which are of degree $n + 1$ in one or the other pair of variables remain. Let

$$L_n(x, y, u, v) \equiv K_n(x, y, u, v) - K_{n-1}(x, y, u, v) \equiv \sum_{i=0}^{n} p_{ni}(x, y)p_{ni}(u, v).$$

Then the aggregate of terms of degree $n + 1$ in u and v in $u \sum_i p_{ni}(x, y)p_{ni}(u, v)$ has the representation

$$\iint_R \rho(r, s)rL_{n+1}(u, v, r, s)L_n(x, y, r, s)dr\, ds,$$

and the terms of degree $n + 1$ in x and y in $x \sum_j p_{nj}(x, y)p_{nj}(u, v)$ are represented by

$$\iint_R \rho(r, s)rL_n(u, v, r, s)L_{n+1}(x, y, r, s)dr\, ds,$$

and as all other terms destroy each other,

$$(u - x)K_n(x, y, u, v) = \iint_R \rho(r, s)rM_n(x, y, u, v, r, s)dr\, ds,$$

where

$$M_n(x, y, u, v, r, s) \equiv L_{n+1}(u, v, r, s)L_n(x, y, r, s) - L_n(u, v, r, s)L_{n+1}(x, y, r, s).$$

Similarly,

$$(v - y)K_n(x, y, u, v) = \iint_R \rho(r, s)sM_n(x, y, u, v, r, s)dr\, ds.$$

Combination of these results gives immediately an identity of corresponding form for

$$[(Au + Bv) - (Ax + By)]K_n(x, y, u, v)$$

with arbitrary A and B. This general identity in a sense takes the place of the Christoffel-Darboux formula for orthogonal systems in one variable.

Similar reasoning is possible in the case of polynomials classified according to the highest exponent attached to either variable. If this exponent is called once more the grade of the polynomial, there are $2n + 1$ normalized orthogonal poly-

nomials of the n-th grade, which may be denoted by $p_{ni}(x, y)$, $i = 0, 1, \cdots, 2n$. In this case $xp_{ni}(x, y)$ may be either of grade $n + 1$ or of grade n, and in the identity

$$xp_{ni}(x, y) = \alpha_{ni}U_{n+1,i}(x, y) + \beta_{ni}V_{ni}(x, y) + \gamma_{ni}W_{n-1,i}(x, y),$$

where the first subscript of U, V, W now indicates the grade of the polynomial in each case, it may be that $\alpha_{ni} = 0$. The relation

$$\alpha_{ni}^2 + \beta_{ni}^2 + \gamma_{ni}^2 \leqq G^2$$

holds as before. The only difference in the form of the identity for

$$(u - x)K_n(x, y, u, v)$$

is that K_n consists now of $(n + 1)^2$ terms and L_n of $2n + 1$ terms,

$$K_n(x, y, u, v) = \sum_{k=0}^{n} \sum_{i=0}^{2k} p_{ki}(x, y)p_{ki}(u, v),$$

$$L_n(x, y, u, v) = \sum_{i=0}^{2n} p_{ni}(x, y)p_{ni}(u, v).$$

Analogous considerations apply also to the theory of orthogonal trigonometric sums in two variables.[10]

UNIVERSITY OF MINNESOTA.

[10] For the case of one variable see the writer's paper *Orthogonal trigonometric sums*, Annals of Mathematics, (2), vol. 34 (1933), pp. 799–814.

ON LOCALLY-CONNECTED AND RELATED SETS

(Second paper)

By S. Lefschetz

The subject matter of three recent papers by the author [1, 2, 3][1] has called forth remarks from Borsuk (on retracts), from Hurewicz (on fixed points) and corrections from Morse (on critical sets), which, together with some further developments induced thereby, we propose to consider in the present paper.

I. Chain-retraction

1. We shall need to refer in the sequel, explicitly and separately, to the following three characteristic properties relating a topological space $\mathfrak{R}$ to its *retract*, the closed set S (Borsuk [6]):

(a) there exists a single-valued transformation $T: \mathfrak{R} \to S$;

(b) T is continuous;

(c) $T = 1$ on S.

As a special case we might have for T a deformation over $\mathfrak{R}$ onto S leaving S point for point invariant. We should then call T a *deformation-retract*.

Once retracts are defined, the notions of AR, ANR follow. We have established in [1] the equivalences between types:

$$\text{ANR} \sim \text{LC}, \tag{1.1}$$

$$\text{AR} \sim \overline{\text{LC}}, \tag{1.2}$$

where $\overline{\text{LC}}$ designates in essence LC sets, in which in addition all spheres are homotopic to points. These two equivalences characterize absolute retracts by properties of local connectedness.

Now one of the chief features of our theory of chain-deformations [2] was the dissociation between the homotopic deformations of a set and its chains, and operations on the chains alone, regardless of what happens to the set itself. The degree to which this was accomplished there did not yield the extension of (1.1), (1.2) to HR sets. In truth we had not looked earnestly for it and were content in our paper to obtain certain other extensions of [1] from LC to HLC. A chance observation by Borsuk, to whom we mentioned this point, led us to the expected generalization as we shall now show. It will be profitable,

Received April 22, 1936.

[1] Numbers in square brackets refer to the bibliography at the end. The general notations and terminology are as in *Topology* [5]; the abridged notations are the same as in [1, 2]: LC = locally connected, H = homology, R = retract, NR = neighborhood retract, A in a compound abridged symbol stands for "absolute".

however, to re-examine the HR notions and indicate how they should be put forth.

2. By referring to [2], No. 9, it will be seen that the HR property there given would be the analogue of (a) if T were a deformation-retract. Furthermore, the analogue of (b) or (c), which would demand, in particular, invariance of the chains on S, was not imposed. In these two deviations may be said to lie the chief difficulty in extending our equivalences. We shall see in fact that when we conform quite strictly with point set retraction, the difficulty vanishes. Agreeing then that the property given, loc. cit., is to be termed chain-shrinking and not described as retraction, we proceed to build up the analogues of point set retraction for chains.

Let then $\mathfrak{K} = \{c_q\}$, $\mathfrak{K}' = \{c'_r\}$ be two quasi-complexes ([2] No. 4). We shall call *chain-transformation* of $\mathfrak{K}$ into $\mathfrak{K}'$ a single-valued transformation τ of the chains of $\mathfrak{K}$ into those of $\mathfrak{K}'$, of form $c_q \to c'_r$, $r \leqq q$, which induces on their chain-groups homomorphisms commutable with the boundary operator F $(F\tau = \tau F)$. It is an *ϵ-transformation* whenever diam $(|c_q| + |\tau c_q|) < \epsilon$ for every c_q of $\mathfrak{K}$. Whenever $\mathfrak{K}'$ is on a set S, we shall also say that $\mathfrak{K}$ is chain-transformed *onto* S. In particular, if τ merely ϵ-transforms a suitable subdivision of $\mathfrak{K}$, we shall still say that $\mathfrak{K}$ is ϵ-chain-transformed onto S.

We now define the closed subset S of $\mathfrak{R}$ as an *homology* retract (= HR) of $\mathfrak{R}$ whenever the following conditions hold:

(α) any finite $\mathfrak{K} = \{c^i_q\}$ of $\mathfrak{R}$ is chain-transformable onto S;

(β) for every ϵ there exists an η such that every $\mathfrak{K}$ of the spherical neighborhood $\mathfrak{S}(S, \eta)$ is ϵ-chain-transformable onto S;

(γ) in both cases (α), (β), the c's $\subset S$ remain invariant.

Whenever (β), (γ) alone hold, S is called a *neighborhood*-HR (= HNR) of $\mathfrak{R}$. This is essentially the extension of the NR notion to the present case.

In the applications to a single chain c_p, we must consider it as the $\mathfrak{K}$ or subchain of the $\mathfrak{K}$ consisting of c_p and $F(c_p)$.

3. The analogy of our three conditions with (a), (b), (c) is obvious. The chief difference is that we do not demand simultaneous transformations of all the chains of $\mathfrak{R}$ onto S.

In the case of an HR one may also have occasion to drop the continuity condition (β). We shall say then that we have a *weak* HR.

A special case of retraction is one in which the transformations are always chain-deformations.[2] We shall say then that we have a deformation HR,

[2] The observation made to us by Borsuk, alluded to at the end of No. 1, consisted precisely in assuming that, everything being immersed in the Hilbert parallelotope $\mathfrak{H}$, the associated deformation-chains be merely taken $\subset \mathfrak{H}$ and not $\subset \mathfrak{R}$. A mild step further consists in frankly replacing chain-deformation by chain-transformation. The latter is intrinsic and may be defined without regard to $\mathfrak{H}$ and in fact, if so desired, for any topological space whatever. The relation between the notion put forth by Borsuk and the

or HNR as the case may be. In particular the first may likewise be weak or otherwise.

The only retraction theorem of [2] is Theorem IX and under our present definitions we must replace in it HNR by "deformation-HNR". We shall refer to the theorem under this form as Theorem IX′ of [2] and similarly for other restated or modified theorems in the sequel.

Let us observe in passing that, as pointed out to us by Morse, the proof of Theorem II of [2], page 9, only establishes the following weaker result, which as a matter of fact, is not really needed in the applications:

THEOREM II′. *If a closed set B is chain-deformed in such a manner that the chains on a closed subset A remain on A, then B may be weakly chain-deformation retracted onto A.*

4. The absolute-HR or -HNR (= AHR, AHNR) are defined as for point set retraction (Borsuk [6]): S is an AHR or AHNR, whenever the HR or HNR property, as the case may be, is possessed by any topological image of S relative to the containing space. It is proved also as for ordinary retraction (Borsuk [6] p. 160) that the necessary and sufficient condition for a compact metric space $\mathfrak{R}$ to be AHR or AHNR is that its image on $\mathfrak{H}$ be HR or HNR for $\mathfrak{H}$ itself. And now we are able to prove that in fact

$$\text{AHNR} \sim \text{HLC}, \tag{4.1}$$

$$\text{AHR} \sim \overline{\text{HLC}}. \tag{4.2}$$

These are the expected analogues of (1.1) and (1.2).

Let us identify $\mathfrak{R}$ throughout with its topological image on $\mathfrak{H}$ and let it first be HLC, with $\xi(\epsilon)$ as its "gauge-function", or the function designated by η in [2], No. 15. Given ϵ and $\eta < \frac{1}{3}\ \xi(\frac{1}{3}\epsilon)$, we shall show that $\mathfrak{R}$ satisfies the HNR conditions with $\eta(\epsilon)$ as the function in condition (β). Let $\mathfrak{K}_p$ be a finite quasi-complex $\subset \mathfrak{S}(\mathfrak{R}, \eta)$ and let mesh $\mathfrak{K}_p < \eta$. If this last condition is not fulfilled, we replace $\mathfrak{K}_p$ by a suitable subdivision of mesh $<\eta$. Let us suppose also that $\mathfrak{K}_p$ has a subcomplex $\mathfrak{L}$ on $\mathfrak{R}$. We define a τ: $\mathfrak{K} \to \mathfrak{R}$, such that $\tau\mathfrak{L} = \mathfrak{L}$ as follows. Take for transform of any c_0 of $\mathfrak{K}_p - \mathfrak{L}$ a zero-chain $c_0' = \tau c_0$ consisting of a point of $\mathfrak{R}$ whose distance to $|c_0| = d(|c_0|, \mathfrak{R})$. Then $\mathfrak{L}$ and the chains c_0' make up a partial realization $\mathfrak{K}_p^*$ of $\mathfrak{K}_p$ whose mesh $< 3\eta < \xi(\frac{1}{3}\epsilon)$. Therefore, by the HLC condition, $\mathfrak{K}_p^*$ may be completed to form an image $\mathfrak{K}_p'$ of $\mathfrak{K}_p$, or transform of $\mathfrak{K}_p$ whose mesh $< \epsilon$; the ϵ-transformation thus determined is precisely a τ of the required nature. This shows that the right side of (4) implies the other.

5. Conversely, suppose that $\mathfrak{R}$ is an AHNR (i.e., an HNR for $\mathfrak{H}$), with $\eta(\epsilon)$ as the function in condition (β). Let $\mathfrak{K}'$ be a partial realization on $\mathfrak{R}$ of a certain

one adopted here may be said to be that whereas he requires some deformation quasi-complex $\mathfrak{D}\mathfrak{K}$, at least in $\mathfrak{H}$, we only demand a partial realization of $\mathfrak{D}\mathfrak{K}$, in the sense of [2], No. 12, with all deformation-chains left out (unrealized).

$\mathfrak{K}$, with mesh $\mathfrak{K}' < \eta(\epsilon)$, where ϵ is given. We first complete $\mathfrak{K}'$ to a realization $\mathfrak{K}''$ of $\mathfrak{K}$ in $\mathfrak{H}$, by the process of [1], No. 17,[3] except that cells are replaced by chains, spheres by chain-boundaries. The sets (ζ), loc. cit., being convex, any cycle on a (ζ) bounds on it. Hence if the boundary $F(c_q)$ of an expected c_q is on (ζ), we may insert $c_q \subset (\zeta)$. This takes the place of the construction by segments, loc. cit.

The complex $\mathfrak{K}''$ thus constructed being of mesh $\leqq$ mesh $\mathfrak{K}' < \eta(\epsilon)$, there exists an ϵ-chain-transformation τ of $\mathfrak{K}''$ into a $\mathfrak{K}^*$ of $\mathfrak{R}$ preserving $\mathfrak{K}'$. It is clear that mesh $\mathfrak{K}^* < \epsilon$ and that it completes $\mathfrak{K}'$ on $\mathfrak{R}$ in the manner prescribed by the HLC condition. Therefore the left side of (4.1) implies the other also and (4.1) is proved.

The same procedure likewise yields the proof of (4.2), since in the two cases now involved the upper bounds of the η's with increasing ϵ are the same as for ϵ.

Corollary. *The sets* AHNR *have the properties of the sets* HLC *considered in* [2], *No.* 16. *In particular, their Betti-groups have the same structure as for a finite complex. The sets* AHR *have the Betti-groups of a point.* (See [1], No. 20.)

It is important to bear in mind throughout that we may have any type of chains for which the postulates of [2], §I, hold. In particular we may have either singular or regular chains (in the sense of that paper).

6. In both [2] and [3] we have had repeated occasion to consider chain shrinking away from a set (see notably [2], p. 16). It was pointed out to us by Morse, however, that the theory of critical sets requires a more delicate notion which we may describe as "local" shrinking away from a set (see Morse [9]). As usual in topology, the term "local" refers to the fact that the given operation may be confined to any preassigned neighborhood of the given set. More precisely, if A, B are subsets of $\mathfrak{R}$, we say that A may be locally chain-shrunk away from B, whenever given any open set $U \supset B$, it is possible to find another V such that $U \supset V \supset B$ and that A may be chain-shrunk onto $A - V$ over AU. By the statement: A may be chain-shrunk away from B at a point x, we shall mean that about x it may be chain-shrunk away from both B and x. That is to say, for every $U \supset x$ there is a $V \supset x$ such that $V \subset U$, and also a neighborhood W of B on U, such that for any quasi-complex $\mathfrak{K}$ there is a chain-deformation displacing only the elements of $\mathfrak{K}$ on U, and this away from both B and x (not bringing them nearer to B or x), those on V being chain-deformed to the outside of $W + V$. Under these definitions we may apply the reasoning of Theorem III of [2], p. 10, without having recourse to Theorem II, the V's of the proof being now as described above, and the W's, loc. cit., being now such that their intersection with the corresponding U plays the rôle of the W

[3] We recall the following errata given at the end of vol. 35 of the Annals of Mathematics and referring to [1], No. 17: line 14 of No. 17, replace "convex sets of $\mathfrak{H}$" by "spheres of $\mathfrak{H}$", line 15, cross out "convex". Cross out line 16.

considered above. We shall designate this stronger theorem as Theorem III′ of [2].

II. Local connectedness and the fixed point formula

7. When we first undertook to extend our basic coincidence and fixed point formulas ([5], Chapter VI) we found that the LC properties of the sets played an important rôle. Confining our attention to the fixed point problem, it was shown that the fixed point formula was valid for a compact LC subset of euclidean spaces ([5], p. 347). Since every finite dimensional space can be mapped topologically onto some euclidean space, this implies the validity of the formula for all finite dimensional compact metric LC spaces. Later we showed ([2], p. 129) that the restriction as to finite dimension could be dropped.

Now the coincidence and fixed point formulas belong in their essence strictly to algebraic topology. One would expect, therefore, to have their range of validity limited, if at all, by restrictions on cycles and the like, that is to say, by HLC rather than LC restrictions. Furthermore, the HLC should refer, preferably, to the more "purely" algebraic homology theories such as those of Vietoris or our own (regular cycles of [2]).

Now in a private communication (Dec. 1935) Hurewicz indicated to us a most ingenious method for establishing *the validity of the fixed point formula for finite dimensional compact metric* HLC *sets in the sense of singular cycles.* Soon after we succeeded in showing that *the formula holds for any compact metric* HLC *set, regardless of type or of dimension.* This result does have the requisite degree of generality, and we shall establish it in the present section. The treatment is independent of the type considered. As a matter of fact, HLC in the sense of singular chains and, say rational coefficients, implies the same for regular or Vietoris chains.

8. Let the general notations be the same as in *Topology*, p. 358, except that the pair $(\mathfrak{H}', L')$ is merely a topological image of $(\mathfrak{H}, L)$. We assume then L to be compact and HLC and shall prove that the basic fixed point formula (49), *Topology*, p. 359, holds for every c.s.v.t. T of L into itself. Or

THEOREM. *Let T be any* c.s.v.t. *of a compact metric* HLC-*set into itself and let φ^p be the matrix of the transformation which T induces on a basis for the rational p-cycles of L. Then the number of signed fixed points of T,*

$$\theta = \Sigma(-1)^p \operatorname{trace} \varphi^p, \tag{8.1}$$

is a topological invariant of T, and if $\theta \neq 0$, T has at least one fixed point.

It will be observed that owing to [2], Theorem VIII, the matrices φ and also the sum in (8.1) are all finite.

Application. If L is an AHR, *every* c.s.v.t. *of T into itself has at least one fixed point.*

Coincidences. While we shall not consider them here, we may remark that the same proof would enable us to show that if L, L' are two HLC spaces and

T, T' are two transformations $L \to L'$ such that T and T'^{-1} are c.s.v.t., then the number θ of their signed coincidences given by formula (48) of [5], p. 359, where all elements are finite, is a topological invariant of the pair T, T'. In particular, when $\theta \neq 0$, there are coincidences of T, T'. Generally speaking, accented elements are to represent the analogues in $\mathfrak{H}'$, but not necessarily the topological images, of the corresponding elements in $\mathfrak{H}$. In particular, the choice of associated pairs N^i, N'^i is to be made as follows. Having determined sequences $\{\mathfrak{N}^i\}$, $\{\mathfrak{N}'^h\}$ converging respectively to L, L' on $\mathfrak{H}$ and $\mathfrak{H}'$, let ϵ_j be the width of $\mathfrak{N}^j$. For a given j we choose an $\mathfrak{N}'^{h_j}$ such that $\mathfrak{N}^j, \mathfrak{N}'^{h_j}$ are related exactly as the associated euclidean space neighborhoods of [5], p. 353. Their projections on $\mathfrak{H}_i, \mathfrak{H}'_i$, i sufficiently high, are precisely N^i, N'^i. These are open neighborhoods of L^i, L'^i whose closures $\bar{N}^i, \bar{N}'^i$ are i-manifolds which may be assumed covered with simplicial complexes of mesh $\eta_i < \epsilon_i$.

Our next step must be the choice of the extensions $\mathfrak{T}_1, \mathfrak{T}_2$, loc. cit. The first is to be the c.s.v.t. $\mathfrak{H} \to \mathfrak{H}'$ image of the identity for $\mathfrak{H}$. The second is to be determined in terms of T as follows. Noticing that as usual we may take L connected, we see that the same will hold for the complex N^i which is then a relative i-circuit. As such it has a fundamental i-cycle Γ_i, whose vertices shall be denoted by x_{ih}. They are of course likewise the vertices of the complex N^i. Now let L'' be the image of T in $L \to L'$. Since T is a c.s.v.t., L'' is homeomorphic to L and hence likewise HLC. Let then y_{ih} be a point of L such that $d(x_{ih}, y_{ih}) = d(x_{ih}, L)$ and let $z_{ih} = Ty_{ih}$, $w_{ih} = y_{ih} \times z_{ih} \subset L''$. The simplexes of N^i make up a quasi-complex $\mathfrak{K}$ of which the set $\{w_{ih}\}$ is a partial realization $\mathfrak{K}'$. Clearly mesh $\mathfrak{K}' \to 0$ with ϵ_i. Hence when it is small enough, $\mathfrak{K}'$ may be completed so as to form a realization $\mathfrak{K}''$ of $\mathfrak{K}$ on L'' whose mesh is less than a certain assigned ξ, and in $\mathfrak{K}''$ the chain Γ_i will have a certain image C_i. The projection of C_i on $N^i \times N'^i$ shall be taken as the component G_i^{2i} defining the term $\mathfrak{T}_{2i}$ of the sequence $\{\mathfrak{T}_{2i}\} = \mathfrak{T}_2$. The rest is then as in [5], p. 358. It is a simple matter to verify that $\mathfrak{T}_2$ is a finite contraction. Therefore the pair $\mathfrak{T}_1, \mathfrak{T}_2$ comes under our conditions of applicability of the fixed point formula. This proves our assertion.

9. The process which enables us to weaken the LC of *Topology* (pp. 347 and 359) into HLC is rather obvious: under the LC assumption we could consider $\mathfrak{K}'$ as a partial realization of the true-complex N^i, then complete it to a true singular image K'' of N^i on L''. K'' is then a true image of a c.s.v.t.: $\bar{N}^i \to L'$ and hence even $\bar{\mathfrak{N}}^i \to L'$. Since all that we needed for our purpose is an i-cycle image of Γ_i, it was sufficient to obtain a partial realization $\mathfrak{K}''$ of the chains, elements of $\mathfrak{K}$, and this the HLC condition enabled us to do.

III. Critical sets

10. Morse has justly criticized certain definitions and results in [3]. From his criticisms it appears that the modifications to be indicated presently must

be made. Except for Theorem XI which must be modified, our proofs are adequate throughout, as they were intuitively based on the proper definitions.

As a preliminary remark it is to be understood that all retractions of [3], or those to be mentioned, are deformation-retractions in the sense of No. 3 of the present paper.

The basic modification required is that throughout [3], with exceptions to be noted presently, chain-shrinking and deformation are to be strengthened by demanding that they be "local" in the sense of No. 6. The exceptions are [3], 4, 6 (from Theorem IV on), 17, where the same must be replaced by chain-deformation retraction.

11. **Analytical and topological critical sets.** The proof of the second part of Theorem XI of [3] ([3], No. 23) rests on the faulty Theorem II of [5] and in fact the theorem is not correct under our definition of t.c.p. This is clearly shown by the function $f = x^3$, the point $x = 0$ being an a.c.p. but not a t.c.p. More generally, an a.c.p. may lack any topological features. This is the type known by experts as *inflectional*. The example just given, however, points also to the necessary modification.

Consider the plane curve $y = x^3$. The reason why the a.c.p. at the origin P is not topological is that the sections of the curve by the lines $y =$ constant are all homeomorphic. On the other hand, if we cut the curve by the parallel pencil $y + mx =$ const., $m \neq 0$ and arbitrarily small, we obtain two ordinary critical points (contacts of ordinary tangents) $\rightarrow P$ when $m \rightarrow 0$. This may be interpreted as follows: the a.c.p. P is the limit of t.c.p.'s of the "modified" function $x^3 + mx$ as $m \rightarrow 0$. Hence Theorem XI is restored provided that we add to t.c.p.'s the points which are limits of t.c.p.'s as $m \rightarrow 0$. This is what we propose to do.

In the case of a function $f(x)$ of $n > 1$ variables a linear increment may be insufficient, notably if the Hessian $H(f) \equiv 0$. This corresponds, for $n = 2$, to surfaces $f = k$ which are developable. We therefore modify f, by adding a quadratic polynomial, into

$$F = f + u_i x_i + \lambda_{ij} x_i x_j, \qquad \lambda_{ij} = \lambda_{ji}. \tag{11.1}$$

We shall now call P a *quasi-topological critical point* (quasi t.c.p.) whenever, given any open set $U \supset P$ on Ω and any $\epsilon > 0$, the parameters u may be chosen $< \epsilon$ in absolute value and such that F has t.c.p.'s in U.

It is now easy to show that Theorem XI holds provided we replace t.c.p. by quasi t.c.p. (Theorem XI′). For this purpose we may assume $\lambda_{ij} = \lambda\delta_{ij}$, and show that λ and the u's may be chosen as required.

Now the reasoning of [3], No. 23, shows that for λ fixed $\neq 0$, $H(F) \not\equiv 0$. Taking $|\lambda| < \epsilon$, $H = 0$ will represent about P an analytical locus of dimension $< n$, and there will be points Q on a given open set $U \supset P$ for which $H \neq 0$. Since $f_{x_i} \rightarrow 0$ when $Q \rightarrow P$, we may find a point Q for which

$$|u_i| = |-(f_{x_i} + \lambda_{ij} x_j)| < \epsilon.$$

Now Q will be an ordinary a.c.p. of F, and hence a t.c.p. of F, on U, with the variables u, λ restricted as required. Therefore P is a quasi t.c.p. and Theorem XI′ follows (compare Morse [10], p. 178).

12. General remark regarding analytical critical points. From the geometric point of view the definition of a.c.p. appears to be too strictly dominated by the analytical features of the basic function $f(x)$. Thus the usual formulation does not cover the case of the function

$$z = \sqrt{1 - x^2 - y^2}, \qquad x^2 + y^2 \leqq 1.$$

To treat this case, it is necessary to consider z as a point-function on an M_2, the sphere. The usual formulation applies even less, of course, to the implicit two-valued function $z(x, y)$ given by

$$x^2 + y^2 + z^2 = 1.$$

Even more striking exceptions could be obtained by taking algebraic surfaces or varieties with singular loci.

From this point of view it would seem more natural to designate as a.c.p. any "new" singular point appearing in the horizontal sections $y = f(x)$ as compared with those below it. This is the analytical analogue of our topological treatment in [3]. This would automatically do away with Whitney's paradoxical arc of critical points, not in a horizontal section [8]. The preceding analytical considerations and similar topological considerations were among our prime notions for attaching the theory of critical sets, not so much to the properties of f as to those of the whole locus[4] $y = f$.

BIBLIOGRAPHY

1. S. Lefschetz, *On locally-connected and related sets*, Annals of Math., vol. 35 (1934), pp. 118–129.
2. ——— *Chain-deformations in topology*, this Journal, vol. 1 (1935), pp. 1–19.
3. ——— *On critical sets*, ibid., pp. 392–412.
4. ——— *Application of chain-deformations to critical points and extremals*, Proc. Nat. Acad. Sci., vol. 21 (1935), pp. 220–222.
5. ——— *Topology*, American Math. Soc. Colloquium Publications, vol. XII, 1930.
6. K. Borsuk, *Sur les retractes*, Fundamenta Math., vol. 17 (1931), pp. 152–170.
7. ———*Über eine Klasse von lokal zusammenhängenden Räumen*, Fundamenta Math., vol. 19 (1932), pp. 220–242.
8. H. Whitney, *A function not constant on a connected set of critical points*, this Journal, vol. 1 (1935), pp. 514–517.
9. M. Morse, *The critical points of a function of n variables*, Trans. Amer. Math. Soc., vol. 33 (1931), pp. 72–91.
10. ——— *The Calculus of Variations in the Large*, American Math. Soc. Colloquium Publications, vol. XVIII, 1934.
11. M. Morse and G. Van Schaack, *Abstract critical sets*, Proc. Nat. Acad. Sci., vol. 21 (1935), pp. 258–263.

Princeton University.

[4] First in our Note [4], then in [3]; see also Morse-van Schaack [11].

THE ALMOST PERIODIC BEHAVIOR OF THE FUNCTION $1/\zeta(1 + it)$

By Aurel Wintner

It is known[1] that the prime-number theorem implies the convergence of the development

$$(1) \qquad 1/\zeta(s) = \sum_1^\infty \mu(n)\, n^{-s},$$

which is obvious in the half-plane $\sigma > 1$, at every point of the line $\sigma = 1$ also. The object of the present note is to show that the *trigonometrical* series

$$(2) \qquad 1/\zeta(1 + it) = \sum_1^\infty \mu(n) n^{-(1+it)} = \sum_1^\infty \mu(n) n^{-1} \exp(-it \log n)$$

is the *Fourier* series of the function which it represents, i.e., that

$$(3) \qquad 1/\zeta(1 + it) \sim \sum_1^\infty \mu(n) n^{-1} \exp(-it \log n),$$

where the sign $\sim$ refers to the class B^2 of Besicovitch.[2] In other words, the function $1/\zeta(1 + it)$ is almost periodic (B^2), and, on placing

$$\mathfrak{M}\{f(t)\} = \lim_{T \to +\infty} \mathfrak{M}_T\{f(t)\},$$

where

$$(4) \qquad \mathfrak{M}_T\{f(t)\} = \int_0^T f(t)dt/T,$$

the mean value

$$(5) \qquad \mathfrak{M}\{e^{i\lambda t}/\zeta(1 + it)\}$$

exists for every real λ and is 0 or $\mu(n)/n$ according as $\lambda \neq \log n$ or $\lambda = \log n$, where $n = 1, 2, \cdots$. On choosing $n = 1$, it follows, in particular, that

$$(6) \qquad \mathfrak{M}\{1/\zeta(1 + it)\}$$

exists and is equal to $\mu(1) = 1$.

Since (3) refers to the class (B^2), it also follows that $\mathfrak{M}\{|\zeta(1 + it)|^{-2}\}$ exists. The latter result, proved by Landau on pp. 801–804 of his *Handbuch*, suggests but does not imply (3); it does not even imply the existence of the Fourier constants (5), (6).

Received December 28, 1935.

[1] Cf. p. 811 of the article by Bohr and Cramér in vol. 2, III_2 of the *Encyklopädie der mathematischen Wissenschaften*, where several references are given.

[2] A. S. Besicovitch, *Almost Periodic Functions*, Cambridge, 1932, Chap. II.

The result is independent of Riemann's hypothesis. On Riemann's hypothesis the almost-periodicity (B^2) of (1) easily follows for $\sigma > \frac{3}{4}$, at least. For on denoting by σ_c the abscissa of convergence of the series (1) and by σ_a the abscissa of its absolute convergence, Littlewood has shown[3] that $\sigma_c = \frac{1}{2}$ on Riemann's hypothesis, while, of course, $\sigma_a = 1$. Thus it is seen from the mean value theorem of Schnee that, on Riemann's hypothesis, $\mathfrak{M}\{| 1/\zeta(\sigma + it) |^{-2}\}$ exists for every $\sigma > \frac{3}{4} = \frac{1}{2}(\sigma_a + \sigma_c)$. Hence, since Littlewood's treatment of the Lindelöf hypothesis implies[4] $1/\zeta(\sigma + it) = O(| t |^{\epsilon})$ uniformly for $\sigma > \frac{1}{2} + \epsilon$, the statement follows from a general result of Besicovitch (op. cit., p. 164).

Let $q > 0$ be a *fixed* integer, $x > q + 1$ a variable which will tend to infinity, and put

$$f_q(t) = \sum_{q+1}^{\infty} \mu(n) n^{-(1+it)} \tag{7}$$

and

$$S_q(t; x) = \sum_{q+1}^{x} \mu(n) n^{-(1+it)}. \tag{8}$$

Thus

$$f_q(t) - S_q(t; x) = \sum_{x+1}^{\infty} \mu(n) n^{-(1+it)}.$$

Hence, on placing $M(x) = \sum_{1}^{x} \mu(n)$ and using the identity

$$n^{-(1+it)} - (n + 1)^{-(1+it)} = (1 + it) \int_n^{n+1} r^{-(2+it)}\, dr,$$

it is seen by partial summation that

$$f_q(t) - S_q(t; x) = (1 + it) \sum_{x+1}^{\infty} M(n) \int_n^{n+1} r^{-(2+it)}\, dr - M(x)[x + 1]^{-(1+it)}.$$

Now[5]

$$M(x) = O(x/\log^5 x) \text{ as } x \to + \infty,$$

so that

$$\sum_{x+1}^{\infty} M(n) \int_n^{n+1} r^{-2}\, dr = \sum_{x+1}^{\infty} O(n/\log^5 n)\, O(n^{-2}) = O(1/\log^4 x)$$

and

$$M(x)[x + 1]^{-1} = O(x/\log^5 x)\, O(x^{-1}) = O(1/\log^4 x).$$

[3] Cf. E. C. Titchmarsh, *The Zeta-function of Riemann*, Cambridge, 1930, p. 78.
[4] Ibid., p. 77.
[5] More than this is known; cf. Bohr and Cramér, loc. cit.

Consequently, if $t > 1$,

(9) $$f_q(t) - S_q(t; x) = tO(1/\log^4 x) \text{ as } x \to +\infty,$$

where the O-term holds uniformly for $1 < t < +\infty$ and q is fixed. Also,[6]

$$1/\zeta(1 - it) = O(\log t) \text{ as } t \to +\infty,$$

while

$$|\bar{f}_q(t) - 1/\zeta(1 - it)| = \left|\sum_1^q \mu(n)n^{-(1-it)}\right| \leqq q = \text{const.}$$

in view of (1) and (7), so that

(10) $$|\bar{f}_q(t)| < C \log t$$

for every $t > e$ and for some $C = C_q > 0$. On combining (9) and (10) with the formal identity

$$|f_q|^2 - |S_q|^2 = -|f_q - S_q|^2 + 2\Re\{(f_q - S_q)\bar{f}_q\},$$

it is seen that

$$|f_q(t)|^2 - |S_q(t; x)|^2 = t^2O(1/\log^8 x) + tO(1/\log^4 x) \log t \text{ as } x \to +\infty,$$

where the O-terms hold uniformly for $e < t < +\infty$. Hence, on using the notation (4),

(11) $$\begin{aligned}\mathfrak{M}_T\{|f_q(t)|^2\} - \mathfrak{M}_T\{|S_q(t; x)|^2\} \\ = O(1/\log^8 x)T^2 + O(1/\log^4 x)\, T \log T \text{ as } x \to +\infty,\end{aligned}$$

where the O-terms hold uniformly for $e < T < +\infty$.

On the other hand, since[7]

$$|S_q(t; x)|^2 - \sum_{q+1}^{x} |\mu(n)/n|^2 = \sum_{q+1}^{x} \sum_{q+1}^{x}{}' \mu(n)\mu(m)(nm)^{-1}(m/n)^{it}$$

in view of (8), it is clear from

$$\int_0^T (m/n)^{it}\, dt = -i\{(m/n)^{iT} - 1\}/\log(m/n), \text{ where } m \neq n,$$

that the absolute value of the difference

(12) $$\mathfrak{M}_T\{|S_q(t; x)|^2\} - \sum_{q+1}^{x} |\mu(n)/n|^2$$

[6] Cf. Titchmarsh, op. cit., p. 17; also p. 24.

[7] The accent in the double summation means that $m \neq n$.

is not greater than

$$T^{-1}\sum_{q+1}^{x}\sum_{q+1}^{x}{}'\,|\,\mu(n)\mu(m)\,(mn)^{-1}\{(m/n)^{iT}-1\}/\log(m/n)\,|$$

$$\leqq T^{-1}\sum_{q+1}^{x}\sum_{q+1}^{x}{}'\,2\,|\,mn\log(m/n)\,|^{-1}.$$

Since q is fixed, the last double sum is[8] $O(\log^2 x)$ as $x\rightarrow+\infty$. Finally, on replacing in (12) the sum $\sum\limits_{q+1}^{x}$ by $\sum\limits_{q+1}^{\infty}$, one commits an error which is but

$$\left|\sum_{x+1}^{\infty}(\mu(n)/n)^2\right|<\sum_{x+1}^{\infty}1/n^2=O(1/x).$$

On combining these estimates of (12) with (11), it follows that

$$\mathfrak{M}_T\{|\,f_q(t)\,|^2\}-\sum_{q+1}^{\infty}|\,\mu(n)/n\,|^2$$
$$=O(1/\log^8 x)T^2+O(1/\log^4 x)T\log T+O(\log^2 x)T^{-1}+O(1/x)\text{ as }x\rightarrow+\infty,$$

where the O-terms hold uniformly for $e<T<+\infty$. On choosing $T=\log^3 x$, the error terms become functions of x alone and tend to zero as $x\rightarrow+\infty$, i.e., as $T\rightarrow+\infty$. Thus if $T\rightarrow+\infty$ and q is fixed, then $\mathfrak{M}_T\{|\,f_q(t)\,|^2\}$ tends to the limit

$$\sum_{q+1}^{\infty}|\,\mu(n)/n\,|^2,$$

so that $\mathfrak{M}\{|\,f_q(t)\,|^2\}$ exists and is equal to the last series. Since this series has a positive value less than $\sum\limits_{q+1}^{\infty}n^{-2}$, it follows that

$$\mathfrak{M}\{|\,f_q(t)\,|^2\}\rightarrow 0\text{ as }q\rightarrow+\infty.$$

This may be written, according to (1) and (7), in the form

$$\mathfrak{M}\left\{\left|\sum_{1}^{q}\mu(n)\,n^{-(1+it)}-1/\zeta(1+it)\right|^2\right\}\rightarrow 0,\ q\rightarrow+\infty.$$

This completes[9] the proof of (3).

THE JOHNS HOPKINS UNIVERSITY.

[8] Cf. Titchmarsh, op. cit., p. 30, where the double sum is shown to be

$$O\left(\sum_{n=1}^{x}n^{1-2\sigma}\sum_{r=1}^{\frac{1}{2}n}r^{-1}\right),$$

where $\sigma=1$, so that the estimate becomes

$$O\left(\sum_{n=1}^{x}n^{-1}\sum_{r=1}^{\frac{1}{2}n}r^{-1}\right)=O\left(\sum_{n=1}^{x}n^{-1}\log n\right)=O\left(\sum_{n=1}^{x}n^{-1}\log x\right)=O\,(\log^2 x).$$

[9] Cf. Besicovitch, op. cit., p. 100 et seq.

MOMENTS OF INERTIA OF CONVEX REGIONS

By Fritz John

Let R denote a closed and bounded two-dimensional convex region. Let d be the greatest, Δ the smallest diameter of R, a diameter being defined as the distance of two parallel lines of support.[1] Let A be the area and L the circumference of R. It was recently proved by F. Behrend that there exist for any R affine transformations transforming R into convex regions for which any one of the following inequalities is satisfied:

$$\frac{d}{\Delta} \leqq \sqrt{2}, \qquad \frac{A}{\Delta^2} \leqq 1, \qquad \frac{d}{L} \leqq \tfrac{1}{4}\sqrt{2};$$

if, moreover, R has a center, i.e., if R is symmetrical with respect to some point, then there are also affine transformations transforming R into regions for which any of the following inequalities hold:

$$\frac{d^2}{A} \leqq 2, \qquad \frac{L^2}{A} \leqq 16, \qquad \frac{L}{\Delta} \leqq 4.$$

The corresponding equalities are all satisfied in the case of a square.

Now let λ denote the ratio of the major and minor axes of the ellipse of inertia of R corresponding to the center of mass of R in a homogeneous mass distribution, i.e., of the "central" ellipse of inertia of R. We shall prove in this paper that the inequalities

$$(1) \qquad \frac{d}{\Delta\lambda} \leqq \sqrt{2} \qquad\qquad (2) \qquad \frac{A}{\Delta^2\lambda} \leqq 1$$

hold; if R has a center, then also

$$(3) \qquad \frac{d^2}{A\lambda} \leqq 2.$$

These inequalities include some of Behrend's results; for every R can be easily transformed by an affine transformation into a region for which the central ellipse of inertia is a circle, i.e., for which $\lambda = 1$, and in this case $d/\Delta \leqq \sqrt{2}$, $A/\Delta^2 \leqq 1$, and if R has a center $d^2/A \leqq 2$ also.

In a second paper I intend to show (1) that if R has a center,

$$\frac{d}{\Delta\lambda} > \frac{\sqrt{2} + \sqrt[3]{100}}{3};$$

Received December 3, 1935.

[1] For notations see *Theorie der konvexen Körper* by Bonnesen and Fenchel; we shall refer to this book as B.-F.

(2) that if B is the area of Legendre's ellipse of inertia of R, then[2]

$$B \leqq \frac{2\pi}{3\sqrt{3}} A.$$

The same methods applied to convex regions in space no longer yield the best possible constants. For example, we obtain for a convex three-dimensional region with a center, for which the central ellipsoid of inertia is a sphere, the inequalities

$$\frac{d}{\Delta} \leqq \sqrt{\frac{10}{3}}, \qquad \frac{V}{\Delta^3} \leqq \frac{\pi}{3}, \tag{4}$$

where V is the volume of the region.

We shall now prove these statements. Let R be a closed and bounded convex region of area A_R. If g is a straight line, we call the distance $D_R(g)$ of the lines of support of R which are orthogonal to g the *diameter of R in the direction of g*.[3] Thus d = Maximum D_R and Δ = Minimum D_R. Let $I_R(g)$ denote the moment of inertia of R with respect to g. If then g and h are any pair of orthogonal straight lines through the center of mass C_R of R, the inequalities

$$\frac{1}{24} A_R D_R^2(h) \leqq I_R(g) \leqq \frac{1}{12} A_R D_R^2(h), \tag{5}$$

$$\frac{A_R^3}{12\, D_R^2(h)} \leqq I_R(h) \tag{6}$$

hold.

In order to prove (6) we do not need the assumption that R is convex. We consider the rectangle Q of the same area A_R which is bounded by the two lines of support of R parallel to g and by two equidistant parallels to h. Its sides are $D_R(h)$ and $A_R/D_R(h)$. Obviously the points contained in Q but not in R have a smaller distance from h than the points contained in R but not in Q. Thus

$$I_R(h) \geqq I_Q(h) = \frac{A_R^3}{12\, D_R^2(h)},$$

and (6) is proved.

We now apply *Steiner's symmetrization*[4] with respect to h; i.e., we carry every secant of R which is normal to h along its straight line, until the center of the secant lies on h. Let $\Sigma_h R = R_1$ denote the region generated from R by sym-

[2] This inequality, together with the more elementary one $B \geqq A$, was proved in a different way by Blaschke, *Über affine Geometrie*, XI, XIV, XIX, in Ber. Verh. d. sächs. Akad. d. Wiss., vols. 69 and 70. Cf. also *Vorlesungen über Differentialgeometrie*, vol. II, p. 64.

[3] In B.-F. called "Breite" of R.

[4] Blaschke, *Kreis und Kugel*, p. 45, or B.-F., p. 69 et seq.

metrization with respect to h. R_1 is again convex and is symmetrical with respect to h. Moreover $A_{R_1} = A_R$, $C_{R_1} = C_R$ (as C_R already lies on h) and

$$D_{R_1}(h) = D_R, \qquad I_{R_1}(g) = I_R(g), \qquad I_{R_1}(h) \leqq I_R(h). \tag{7}$$

h is a double normal of R_1,[5] i.e., the lines perpendicular to h in the points of intersection of h with the boundary $\overline{R_1}$ of R_1 are lines of support of R_1.

Now let $R_2 = \Sigma_g R_1$. Then R_2 will be symmetrical with respect to g and h. Moreover,

$$I_{R_2}(g) \leqq I_{R_1}(g), \qquad D_{R_2}(h) = D_{R_1}(h), \tag{8}$$

since h is a double normal of R_1. We consider the rhomb S of area $A_{R_2} = A_R$ which has as two opposite vertices the points of intersection of h with the boundary of R_2. *Because of the convexity* and symmetry of R_2 the distance of any point of R_2 not belonging to S from g is greater than that of any point of S not belonging to R_2. Thus

$$I_{R_2}(g) \geqq I_S(g) = \frac{1}{24} A_{R_2} D^2_{R_2}(h).$$

From this relation, together with (7) and (8), it follows that

$$I_R(g) \geqq \frac{1}{24} A_R D^2_R(h).$$

We now return to R_1. Let g^* be the perpendicular bisector of the secant h of R_1. As g passes through the center of mass of R_1 and is parallel to g^* we have

$$I_{R_1}(g) \leqq I_{R_1}(g^*). \tag{9}$$

We reflect that part of R_1 which lies on one side of h with respect to g^*, leaving the rest of R_1 unchanged. As R_1 is symmetrical with respect to h, the result will be a convex region R_3 symmetrical with respect to the point of intersection P of g^* and h. Moreover,

$$D_{R_1}(h) = D_{R_3}(h), \quad A_{R_1} = A_{R_3}, \quad I_{R_1}(g^*) = I_{R_3}(g^*). \tag{10}$$

Finally let $R_4 = \Sigma_h R_3$. R_4 is symmetrical with respect to h and P and therefore with respect to g^*. Let Q be the rectangle of area $A_{R_4} = A_R$, which is bounded by the lines of support of R_4 parallel to g^* and by two equidistant parallels to h. Then, since R_4 is convex and symmetrical with respect to g^* and h,

$$I_{R_3}(g^*) = I_{R_4}(g^*) \leqq I_Q(g^*) = \frac{1}{12} A_{R_4} D^2_{R_4}(h). \tag{11}$$

[5] See B.-F., p. 52.

From (7), (9), (10), (11) we may then conclude

$$I_R(g) \leqq \frac{1}{12} A_R D_R^2(h).$$

Thus (5) is proved.

From (5) it follows that

$$\frac{d^2}{\Delta^2} = \frac{\text{Maximum } D_R^2}{\text{Minimum } D_R^2} \leqq 2 \frac{\text{Maximum } I_R}{\text{Minimum } I_R} = 2\lambda^2,$$

according to the definition of the ellipse of inertia.

Moreover, we get from (5) and (6)

$$\frac{A_R^3}{12\, D_R^2(h)} \leqq I_R(h), \qquad I_R(g) \leqq \frac{1}{12} A_R D_R^2(h),$$

or

$$\lambda^2 \geqq \frac{I_R(h)}{I_R(g)} \geqq \frac{A_R^2}{D_R^4(h)},$$

or

$$A_R \leqq \lambda\, D_R^2(h)$$

for every h. In particular,

$$A_R \leqq \lambda\Delta^2.$$

Now let R have a center (i.e., R is symmetrical with respect to C_R). Let $D_R(h)$ be a greatest diameter: $D_R(h) = d$. Then h will be a double normal of R.[6] Let $R_2 = \Sigma_g R$. R_2 will be symmetrical with respect to g and C_R and therefore with respect to h. Moreover, $A_{R_2} = A_R$, $I_{R_2}(h) = I_R(h)$, $D_{R_2}(h) = D_R(h) = d$. We consider the same rhomb S as we did after (8); from the convexity and symmetry of R_2 it follows immediately that

$$I_{R_2}(h) \leqq I_S(h) = \frac{1}{6} \frac{A_R^3}{D_{R_2}^2(h)} = \frac{1}{6} \frac{A_R^3}{d^2}.$$

As according to (5)

$$\frac{1}{24} A_R d^2 = \frac{1}{24} A_R D_R^2(h) \leqq I_R(g),$$

it follows that

$$\frac{A_R d^2}{24} \leqq \frac{I_R(g)}{I_R(h)} \cdot \frac{A_R^3}{6d^2} \leqq \lambda^2 \frac{A_R^3}{6d^2}.$$

Thus $d^2 \leqq 2\lambda A_R$, if R has a center.

In the case of a three-dimensional convex solid B with a center C we may

[6] See B.-F., p. 52.

proceed in the same manner. Let H denote a plane through C; let g be the normal of H in C. $I_B(H)$ and $I_B(g)$ may denote the moments of inertia of B with respect to H and g respectively, and V the volume of B. Let $D_B(g)$ be the distance of the planes of support parallel to H.

Then $I_B(H)$, $D_B(g)$ and V are unaltered, and $I_B(g)$ is diminished, if we apply to B symmetrization with respect to some plane containing g. By a sequence of such symmetrizations one can transform B into a solid of revolution B_1 with axis g. This is the *construction of Schwarz*,[7] which consists in replacing every plane section of B parallel to H by a circle of the same area and with center on g. B_1 is again a convex solid (Theorem of Brunn) of the same volume V and is symmetrical with respect to H. Besides

$$I_{B_1}(H) = I_B(H), \qquad I_{B_1}(g) \leqq I_B(g), \qquad D_{B_1}(g) = D_B(g).$$

If the Q is the cylinder of revolution with axis g and volume V which is bounded by the two planes of support of B_1 parallel to H, then

$$I_{B_1}(g) \geqq I_Q(g) = \frac{V^2}{2\pi D_{B_1}(g)},$$

i.e.,

$$I_B(g) \geqq \frac{V^2}{2\pi D_B(g)}. \tag{12}$$

Moreover, as B_1 is a convex solid of revolution and symmetrical with respect to H, we have

$$I_{B_1}(H) \leqq I_Q(H) = \frac{1}{12} V\, D^2_{B_1}(g);$$

thus

$$I_B(H) \leqq \frac{1}{12} V\, D^2_B(g). \tag{13}$$

Finally, we compare B_1 with the double cone S of volume V which is symmetrical with respect to H and the vertices of which are the points of intersection of g with the planes of support orthogonal to g. Then

$$I_{B_1}(H) \geqq I_S(H) = \frac{1}{40} V\, D^2_{B_1}(g),$$

i.e.,

$$I_B(H) \geqq \frac{1}{40} V\, D^2_B(g). \tag{14}$$

In the particular case where the central ellipsoid of inertia is a sphere, we have

$$I_B(g) = 2I_B(H) = \text{const.}$$

[7] Blaschke, *Kreis und Kugel*, p. 86, or B.-F., pp. 71–72.

Thus it follows from (13) and (14) that

$$\frac{d}{\Delta} = \frac{\text{Maximum } D_B}{\text{Minimum } D_B} \leqq \sqrt{\frac{10}{3}},$$

and from (12) and (13) that

$$V \leqq \frac{\pi}{3} D_B^3(g)$$

for every g, i.e.,

$$V \leqq \frac{\pi}{3} \Delta^3.$$

UNIVERSITY OF KENTUCKY.

BLOCH FUNCTIONS

By Raphael M. Robinson

In this paper we prove the following theorem.

If $f(x) = x + \cdots$ is regular in $|x| < 1$, and maps $|x| < 1$ on a (many-sheeted) region such that the upper bound of the radii of circles contained in a single sheet of the region is as small as possible, then the unit circle is a natural boundary for $f(x)$.

In proving this, we introduce a method which can probably be used to obtain much more extended results about the functions which map $|x| < 1$ on regions not containing circles any larger than necessary.

This paper is divided into three sections. §1 contains some preliminary material concerning Bloch's Theorem. §2 contains some lemmas about special mapping functions. §3 contains the above theorem and another similar theorem, and some remarks concerning further results about the functions mentioned above.

1. Let R be a region in the complex plane, and let $f(x)$ be regular in R. Then

(1) $f(x)$ is said to be univalent (= schlicht) in R, if $f(x_1) \neq f(x_2)$ for $x_1 \neq x_2$, x_1 and x_2 in R.

(2) If S is a point set in the complex plane, $f(x)$ is said to assume S in R, if for every y in S there is an x in R, such that $f(x) = y$.

(3) If S is a point set in the complex plane, $f(x)$ is said to assume S univalently in R, if $f(x)$ is univalent in a subregion R_1 of R, and assumes S in R_1.

Bloch's theorem may be stated in the following form.

If $f(x) = x + \cdots$ is regular in $|x| < 1$, there is a complex number y_0 such that $f(x)$ assumes $|y - y_0| < P$ univalently, where $P > 0$ is an absolute constant.

Here y_0 depends on the function $f(x)$, but P does not.

There are three constants connected with this theorem defined as follows. $\mathfrak{B}$ is the upper bound of constants P which satisfy the theorem. $\mathfrak{L}$ is the upper bound of P if we strike out the word "univalently". $\mathfrak{A}$ is the upper bound of P if $f(x)$ is assumed to be univalent; here it is immaterial whether the word "univalently" is present or not.

The relations $\mathfrak{B} \leqq \mathfrak{L} \leqq \mathfrak{A}$ are obvious. Landau[1] has given numerical bounds for the values of the three constants; in particular, he has shown that $\mathfrak{L} < \mathfrak{A}$. An improved upper bound for $\mathfrak{A}$ was given by me.[2]

Received February 3, 1936.

[1] E. Landau, *Über die Blochsche Konstante* ···, Math. Zeitschrift, vol. 30 (1929), pp. 608–634.

[2] Robinson, *The Bloch Constant* $\mathfrak{A}$ ···, Bull. Amer. Math. Soc., vol. 41 (1935), pp. 535–540.

Let $\mathfrak{C} = \mathfrak{B}, \mathfrak{L}$, or $\mathfrak{A}$, with the following understanding. When $\mathfrak{C} = \mathfrak{B}$, "assume properly" shall mean "assume univalently"; otherwise, it shall mean simply "assume". By an admissible function we shall mean a function $f(x) = x + \cdots$ which is regular in $|x| < 1$, and, if $\mathfrak{C} = \mathfrak{A}$, is univalent in $|x| < 1$.

Then the definitions of $\mathfrak{B}$, $\mathfrak{L}$, and $\mathfrak{A}$ go into the following statement. $\mathfrak{C}$ is the upper bound of constants P such that every admissible function $f(x)$ assumes the interior of a circle of radius P properly.

If $\epsilon > 0$, not every admissible function $f(x)$ properly assumes the interior of a circle of radius $\mathfrak{C} + \epsilon$. We wish to show that there are admissible functions which do not properly assume the interior of any circle of radius greater than $\mathfrak{C}$.

For every positive integer n, choose an admissible function $g_n(x)$ which does not properly assume the interior of any circle of radius $\mathfrak{C} + 1/n$, and which satisfies the inequality[3]

$$|g_n'(x)| \leqq \frac{1}{1 - |x|^2} \qquad \text{for } |x| < 1.$$

Then the sequence $g_n(x)$ is uniformly bounded in the circle $|x| \leqq r$, if $r < 1$. Hence a subsequence can be chosen which converges uniformly in any circle $|x| \leqq r$, $r < 1$, and hence converges to an analytic function $f(x)$ in $|x| < 1$.

We thus obtain a sequence $f_n(x)$ such that $f_n(x)$ is admissible, does not properly assume the interior of a circle of radius $\mathfrak{C} + 1/n$, and such that $f_n(x)$ converges uniformly to $f(x)$ in $|x| \leqq r$, for every $r < 1$.

We wish to show that $f(x)$ is an admissible function which does not properly assume the interior of any circle of radius greater than $\mathfrak{C}$. In the first place, $f(x)$ is admissible. For $f(x)$ is regular in $|x| < 1$, of the form $x + \cdots$, and if the $f_n(x)$ are univalent, so also is $f(x)$. It remains to be shown that $f(x)$ does not properly assume the interior of any circle of radius $> \mathfrak{C}$. We shall distinguish two cases, according as "assume properly" means (1) "assume" or (2) "assume univalently".

(1) Suppose that, in $|x| < 1$, $f(x)$ assumes the circle $|y - y_0| < \mathfrak{C} + 2\epsilon$, where y_0 is a complex number, and $\epsilon > 0$. To any point y_1 in $|y - y_0| \leqq \mathfrak{C} + \epsilon$ there is an $r_1 > 0$, and a function $h_1(y)$ regular in $|y - y_1| \leqq 2r_1$, except possibly for a branch point of finite order at y_1, and such that in $|y - y_1| \leqq 2r_1$, $|h_1(y)| < 1$ and $f(h_1(y)) = y$ (i.e., $h_1(y)$ is a branch of the function inverse to $f(x)$). We can find a finite number of points $y_1, \cdots, y_m$, and corresponding $r_1, \cdots, r_m$, such that the circles $|y - y_k| < r_k$ $(1 \leqq k \leqq m)$ cover $|y - y_0| \leqq \mathfrak{C} + \epsilon$. The values assumed by the function $h_k(y)$ (corresponding to the point y_k) in $|y - y_k| < 2r_k$ form a region R_k in the interior of the unit circle. $f(x)$ maps R_k on the circle $|y - y_k| < 2r_k$, possibly counted multiply. Since $f_n(x)$ converges to $f(x)$ uniformly in R_k, $f_n(x)$ maps R_k on a region containing $|y - y_k| < r_k$, for $n > N_k$. Hence for $n > N$ $(= \max N_k)$, $f_n(x)$ assumes $|y - y_0| < \mathfrak{C} + \epsilon$.

[3] Landau, loc. cit., p. 618.

(2) Suppose that, in $|x| < 1$, $f(x)$ assumes the circle $|y - y_0| < \mathfrak{C} + 4\epsilon$ univalently, where y_0 is a complex number, and $\epsilon > 0$. There is a region R_4 in $|x| < 1$, such that $f(x)$ is univalent in R_4 and maps R_4 on $|y - y_0| < \mathfrak{C} + 4\epsilon$. There are regions R_2 and R_3 within R_4, which $f(x)$ maps on $|y - y_0| < \mathfrak{C} + 2\epsilon$ and $|y - y_0| < \mathfrak{C} + 3\epsilon$, respectively. In R_3, $f(x)$ is univalent, and $f_n(x)$ converges uniformly to $f(x)$; hence for $n > N_1$, $f_n(x)$ is univalent in R_2. For $n > N\ (\geqq N_1)$, $f_n(x)$ assumes $|y - y_0| < \mathfrak{C} + \epsilon$ in R_2.

Hence (in either case), if $f(x)$ properly assumes the interior of a circle of radius $\mathfrak{C} + 4\epsilon$, $f_n(x)$ properly assumes a circle of radius $\mathfrak{C} + \epsilon$, for $n > N$; but since $f_n(x)$ does not properly assume a circle of radius $\mathfrak{C} + 1/n$, this is impossible. Therefore $f(x)$ does not assume a circle of radius $> \mathfrak{C}$ properly.

Any admissible function which does not assume the interior of a circle of radius greater than $\mathfrak{C}$ properly will be called a Bloch function. It will be said to be of the first, second, or third kind, according as $\mathfrak{C} = \mathfrak{B}$, $\mathfrak{L}$, or $\mathfrak{A}$.

We have shown that there are Bloch functions of each kind. If $f(x)$ is a Bloch function, so also is $f(\alpha x)/\alpha$, where $|\alpha| = 1$; it is not known whether Bloch functions are unique except for this trivial transformation. Bloch functions of the first and second kind are not univalent, since $\mathfrak{B}$ and $\mathfrak{L}$ are less than $\mathfrak{A}$; it is not known whether Bloch functions of the first and second kinds are different from each other.

2. By the inner radius of a region R with respect to a point a in R is meant the number $\rho > 0$ such that $|x| < \rho$ can be mapped on R by a function of the form $a + x + \cdots$. An equivalent statement is that $|x| < 1$ can be mapped on R by a function of the form $a + \rho x + \cdots$.

LEMMA 1. *Let $0 < \theta < \pi$, $1 \leqq \alpha \leqq 2$. Let $R(\alpha, \theta)$ be the region obtained from $|y| < 1$ by modifying the boundary in the following way. Replace the arc of the unit circle joining $e^{\pm i\theta}$, and passing through 1, by another circular arc joining $e^{\pm i\theta}$, and such that the angle formed at $e^{\pm i\theta}$, measured within R, is $\alpha\pi$. Then the inner radius of $R(\alpha, \theta)$ with respect to the origin is*

$$\frac{\alpha \sin(\theta/\alpha)}{\sin\theta}.$$

Proof. If $\mathfrak{J}w > 0$, we may write $w = |w|e^{i\varphi}$, where $0 < \varphi < \pi$; by w^α we shall mean $|w|^\alpha e^{i\alpha\varphi}$. The transformation

$$\frac{y - e^{i\theta}}{e^{i\theta}y - 1} = \left(\frac{z - e^{i\theta}}{e^{i\theta}z - 1}\right)^\alpha$$

takes $|z| < 1$ into $R(\alpha, \theta)$. The point $y = 0$ corresponds to a point z determined by $e^{i\theta/\alpha} = (z - e^{i\theta})/(e^{i\theta}z - 1)$. This gives $z = -a$, where

$$a = \frac{e^{i\theta} - e^{i\theta/\alpha}}{e^{i\theta}e^{i\theta/\alpha} - 1} = \frac{\sin[(\theta - \theta/\alpha)/2]}{\sin[(\theta + \theta/\alpha)/2]},$$

so that $a \geqq 0$. If we put

$$z = \frac{u - a}{1 - au},$$

which takes $|u| < 1$ into $|z| < 1$, then $u = 0$ corresponds to $z = -a$, and hence to $y = 0$.

Now

$$\left(\frac{dz}{du}\right)_{u=0} = 1 - a^2, \qquad \left(\frac{dy}{dz}\right)_{z=-a} = \frac{\alpha(e^{i\theta/\alpha})^{\alpha-1}}{(e^{i\theta}a + 1)^2}.$$

Multiplying these together, substituting the value of a, and simplifying, we have

$$\left(\frac{dy}{du}\right)_{u=0} = \frac{\alpha \sin (\theta/\alpha)}{\sin \theta}.$$

Since the transformation from u to y takes $|u| < 1$ into $R(\alpha, \theta)$, in such a way that $u = 0$ goes into $y = 0$, this is the required inner radius.

Lemma 2. *The inner radius, with respect to the origin, of the region obtained from $|u| < 1$ by removing the points $r \leqq u < 1$, where $0 < r \leqq 1$, is*

$$p = \frac{4r}{(1 + r)^2}.$$

Proof. Since $v = x/(1 + x)^2$ maps $|x| < 1$ on the v-plane excluding the half-line $v \geqq 1/4$, the required mapping function is determined from

$$\frac{u}{(1 + u)^2} = p\frac{x}{(1 + x)^2},$$

with the condition $|u| < 1$, and this transformation is of the form

$$u = px + \cdots .$$

Lemma 3. *The inner radius, with respect to the origin, of the region obtained from $R(\alpha, \theta)$ by excluding the points for which $y \geqq 1 - b\theta$ is greater than 1, if $1 < \alpha < 2$,*

$$b < \cot (\alpha \operatorname{arc} \cot [2(\alpha^2 - 1)/3]^{1/2}) = \beta,$$

and θ is sufficiently small. The number β cannot be replaced by any larger number.

Proof. The circle $|x| < 1$ can be mapped on the region obtained from $|u| < 1$ by excluding the points $r \leqq u < 1$, by a transformation of the form $u = px + \cdots$. $|u| < 1$ can be mapped on $R(\alpha, \theta)$ by a transformation of the form $y = ([\alpha \sin (\theta/\alpha)]/\sin \theta)u + \cdots$. Combining these transformations, we map $|x| < 1$ on a region obtained from $R(\alpha, \theta)$ by excluding certain points on the real axis. We can choose p so that $y = x + \cdots$ (i.e., so that the inner radius of the last region with respect to the origin is 1). Suppose that in this case the end of the cut is at the point $1 - \kappa\theta$, where κ depends on θ and α. The proof will be complete if we can show that $\kappa \to \beta$ as $\theta \to 0$. This in turn

will follow if the position of the end of the cut is given by a power series in θ of the form $1 - \beta\theta + \cdots$.

Choosing the value of p stated above, we have

$$p = \frac{\sin\theta}{\alpha\sin(\theta/\alpha)} = 1 - \frac{1}{6}\left(1 - \frac{1}{\alpha^2}\right)\theta^2 + \cdots,$$

the series containing only even powers of θ. If r has the meaning of Lemma 2, then it can be calculated from p, and the result is

$$r = 1 - 2c\theta + 2c^2\theta^2 + \cdots,$$

where

$$c = [(1 - 1/\alpha^2)/6]^{1/2}.$$

Let a have the meaning of Lemma 1. Then a simple calculation shows that

$$a = \frac{\alpha - 1}{\alpha + 1}\left(1 + \frac{\theta^2}{6\alpha} + \cdots\right).$$

The point $x_0 = 1$ goes into $u_0 = r$. The transformation used in Lemma 1 takes this into the point $z_0 = (r - a)/(1 - ar)$. Putting in the values of r and a, and simplifying, we find that

$$z_0 = 1 - 2\alpha c\theta + 2\alpha^2c^2\theta^2 + \cdots.$$

With this value of z_0, we find that

$$\frac{z_0 - e^{i\theta}}{e^{i\theta}z_0 - 1} = \frac{2\alpha c + i}{2\alpha c - i}(1 + A\theta^2 + \cdots),$$

where A depends on α in a manner which we do not need to determine. Hence, for the corresponding point y_0,

$$\frac{y_0 - e^{i\theta}}{e^{i\theta}y_0 - 1} = \lambda(1 + \alpha A\theta^2 + \cdots),$$

where

$$\lambda = \left(\frac{2\alpha c + i}{2\alpha c - i}\right)^{\alpha}.$$

Solving this for y_0, we find

$$y_0 = 1 - \frac{\lambda + 1}{\lambda - 1}i\theta + \cdots.$$

Now $(2\alpha c + i)/(2\alpha c - i)$ is a quantity whose absolute value is 1, and whose argument is 2 arc cot $2\alpha c$; hence

$$\lambda = e^{2i\gamma},$$

where

$$\gamma = \alpha \operatorname{arc\,cot} 2\alpha c.$$

Therefore

$$\frac{\lambda + 1}{\lambda - 1} i = \frac{e^{i\gamma} + e^{-i\gamma}}{e^{i\gamma} - e^{-i\gamma}} i = \cot \gamma.$$

But

$$\cot \gamma = \cot (\alpha \text{ arc } \cot 2\alpha c) = \cot (\alpha \text{ arc } \cot [2(\alpha^2 - 1)/3]^{1/2}) = \beta,$$

so that

$$y_0 = 1 - \beta\theta + \cdots .$$

Thus the position of the end of the cut is expanded in a series of the required form.

LEMMA 4. *The inner radius with respect to the origin of the region $R_1(\theta)$ obtained from $| y | < 1$ by adding the interior of the circle through $e^{i\theta}$ and $e^{-i\theta}$ orthogonal to the unit circle, and then taking away the points for which $y \geqq 1 - \theta/3$, is greater than 1, if $0 < \theta < \theta_0$, where θ_0 is an absolute constant.*

Proof. This is a special case of Lemma 3, since

$$1/3 < \cot [(3/2) \text{ arc } \cot (5/6)^{1/2}] = 0.336 \cdots .$$

3. The theorem stated at the beginning of the paper may now be easily proved.

THEOREM 1. *Every Bloch function of the first kind has the unit circle as a natural boundary.*

Proof. Suppose that the unit circle is not a natural boundary for every Bloch function of the first kind. Then we can find a Bloch function $f(x)$ of the first kind which is regular at 1. There is a positive $\delta < \theta_0$ such that $f(x)$ is regular in the region $R_1(\theta)$ of Lemma 4 for every $\theta < \delta$, and maps $R_1(\theta)$ on a region which is obtained from the map of $| x | < 1$ by replacing an analytic boundary arc J, which is approximately a straight line, by a new analytic arc K, which is approximately a semicircle on J as a diameter, and which lies on the outside of the map of $| x | < 1$, and then making a cut L along an analytic arc which is approximately a diameter of K, orthogonal to J, the length of the cut L being approximately two thirds of the diameter of K. When we say that this part of the boundary is approximately a certain shape, we mean that it can be made as nearly that as we please, by taking θ sufficiently small. Since $R_1(\theta)$ has an inner radius greater than 1 with respect to the origin, its map must contain circles of radius greater than $\mathfrak{B}$. But it is clear that, if θ is sufficiently small, replacing the boundary arc J of the map of $| x | < 1$ by K and L could not increase the size of the circles in the map.

THEOREM 2. *Let $f(x)$ be a Bloch function of the third kind, and map $| x | < 1$ on a region R. If an analytic arc D is part of the boundary of R, there are points of R on each side of D in the neighborhood of any point of D.*

Proof. D is the map of an arc E of the unit circle along which $f(x)$ is regular and $f'(x) \neq 0$. Without loss of generality, let 1 be a point of E, and $f(1)$ the

point of D under consideration. Then there is a positive $\delta < \theta_0$, such that $f(x)$ is regular in the region $R_1(\theta)$ of Lemma 4 for every $\theta < \delta$. If R lies only on one side of D in the neighborhood of $f(1)$, then $R_1(\theta)$ is mapped by $f(x)$ on a region which does not overlap itself, if θ is sufficiently small. The further argument is then as in Theorem 1.

Suppose that $f_1(x)$ is a Bloch function of the first kind, and maps $|x| < 1$ on a many-sheeted region R_1. We should like to show that R_1 has no boundary. (Branch points of any order in R_1 would serve to limit the size of the circles assumed univalently by $f_1(x)$, provided that these branch points were suitably distributed.) R_1 cannot have an analytic boundary arc, as Theorem 1 shows. On the other hand, R_1 cannot have a boundary arc which is not sufficiently smooth to allow a circle of radius $\mathfrak{B}$ to roll along it, on the inside of R_1. For in this case we could modify R_1 by simply adding something to it, in such a way as to increase its inner radius with respect to the origin, but not the size of the circles in it. The remaining case is intermediate between these. In this case, one might try to use a method similar to that used in the case of an analytic arc; but in this case the modification of the boundary would have to be made directly in the map, instead of first in the unit circle. The difficulty comes in calculating the effect of such a modification of the boundary of R_1 on the inner radius of R_1 with respect to the origin. Were this difficulty overcome, and the results as anticipated, the same method could be used to show that a Bloch function $f_2(x)$ of the second kind maps $|x| < 1$ on a boundaryless many-sheeted region R_2 (branch points of infinite order in R_2 serving to limit the size of the circles assumed by $f_2(x)$), and that a Bloch function $f_3(x)$ of the third kind maps $|x| < 1$ on a plane region R_3 of which every point in the plane is an interior or boundary point.

BROWN UNIVERSITY.

AN EXTENSION OF THE TABLE OF BERNOULLI NUMBERS

By D. H. Lehmer

As part of an investigation on Fermat's Last Theorem being conducted by Professor H. S. Vandiver under the auspices of the American Philosophical Society, it was thought advisable to extend the existing tables of Bernoulli numbers to be used for direct divisibility tests in seeking to establish the regularity[1] of primes. Congruence properties of Bernoulli numbers are important in other branches of the theory of numbers such as class-number problems, and it is hoped that the table given below will prove useful in these and other connections.

Previous tables of the numbers of Bernoulli have been given by Euler,[2] Ohm,[3] Adams[4] and Serebrennikoff.[5] The table of Adams gives the first 62 (non-zero) numbers of Bernoulli, while Serebrennikoff's calculations extend to the first 92 numbers, the first 90 of which have been reprinted by J. Peters[6] and H. T. Davis.[7] The present table gives 20 additional numbers, thus making available the first 110 Bernoulli numbers.

The method used by Adams and Serebrennikoff (except for his last number) was based on the fundamental though inefficient umbral recurrence relation

$$(B + 1)^n = B^n \qquad (n > 1), \tag{1}$$

the fractional terms being eliminated by means of the von Staudt-Clausen theorem.[4] A recurrence with fewer terms than (1) would have been preferable. Many such recurrences involving either B_n or numbers closely allied to B_n are available.[8] As the recurrences become shorter, however, the coefficients ultimately increase in complexity. The best compromise[9] seems to be the following lacunary recurrence for the so-called numbers of Genocchi,[10] which occur in the expansion of $\tan(x/2)$ and are connected with the Bernoulli numbers by

$$G_n = 2(1 - 2^n)B_n. \tag{2}$$

Received January 29, 1936.

[1] A prime p is said to be regular if it does not divide the numerator of any of the first $(p - 3)/2$ non-zero Bernoulli numbers.

[2] Euler, Acta Petropolitanæ, vol. 5 (1781), Part 2, p. 46.

[3] Ohm, Jour. für Math., vol. 20 (1840), p. 11. (Table computed by Rothe.)

[4] Adams, Jour. für Math., vol. 85 (1878), pp. 269–272; Cambridge Observations, vol. 22, appendix; British Association Reports, Sectional Transactions, 1877, pp. 10–11; *Collected Papers*, vol. 1, 1896, pp. 425–451.

[5] Serebrennikoff, Academia Nauk, (8), vol. 16 (1905), No. 10; ibid., vol. 19 (1906), No. 4.

[6] J. Peters, *Zehnstellige Logarithmentafel*, vol. 1, Berlin, 1922, table 8, p. 83.

[7] H. T. Davis, *Tables of Higher Mathematical Functions*, vol. 2, Bloomington, 1935, pp. 230–233.

[8] See N. Nielsen, *Traité Elémentaire des Nombres de Bernoulli*, Paris, 1923, for references.

[9] Other methods for computing large Bernoulli numbers such as that used by Serebrennikoff in computing N_{92} and the central difference method of Joffe [Quarterly Jour. of Math., vol. 47 (1916), pp. 103–126, and vol. 48 (1917–20), pp. 193–271] were also considered.

[10] Genocchi, Jour. für Math., vol. 99 (1886), pp. 315–316.

The recurrence[11] in question is

$$4G_{2n} + 3\sum_{\lambda=1}^{[n/3]} G_{2n-6\lambda}\binom{2n}{6\lambda} = \begin{cases} 2n, \text{ if } n = 3K - 1, \\ -4n, \text{ otherwise.} \end{cases} \tag{3}$$

As it stands, this is not in the form best suited for calculation. If we define

$$g(n, \lambda) = \binom{2n}{6\lambda} \mid G_{2n-6\lambda} \mid, \tag{4}$$

(3) becomes

$$100 \mid G_{2n} \mid = 75 \sum_{\lambda=1}^{[n/3]} (-1)^{\lambda-1} g(n, \lambda) + \begin{cases} 50\, n(-1)^{n} \\ 100\, n(-1)^{n-1}. \end{cases} \tag{5}$$

For $\lambda > 0$, the $g(n, \lambda)$ are readily computed from previous g's by

$$g(n, \lambda) = g(n - 3, \lambda - 1) f(n, \lambda),$$

where

$$f(n, \lambda) = \frac{2n(2n - 1) \cdots (2n - 5)}{6\lambda(6\lambda - 1) \cdots (6\lambda - 5)}.$$

Since $g(n, \lambda)$ is an integer, the denominator of $f(n, \lambda)$, when it is in its lowest terms, must divide $g(n - 3, \lambda - 1)$. This serves as a useful check in computing $g(n, \lambda)$. The disadvantages in handling the somewhat larger numbers G_n are more than offset by two advantages: (A) the G's are integers, so that no additional work in eliminating fractions is necessary, and (B) an excellent check is afforded by writing

$$\mid B_{2n} \mid = N_n / D_n, \tag{6}$$

and thus obtaining from (2)

$$N_n = G_{2n} / d_n, \tag{7}$$

where d_n is the integer $2(4^n - 1)/D_n$, and thus obtaining N_n as the quotient of an exact large-number division. (For $n = 110$, G_{220} has 313 digits, while d_{110} has 63 digits.) This check assures the correctness not only of the present table but also of the tables of Adams and Serebrennikoff inasmuch as their values were used in computing the table[12] of G_n.

The arrangement of the numbers in the table of Bernoulli numerators is a departure from the traditional method of writing large numbers in a few long lines. It will be found that with standard computing machines large numbers may be dealt with more effectively if grouped in "periods" of 9 figures and written in columns.

[11] D. H. Lehmer, Annals of Math., vol. 36 (1935), pp. 637–649. The N_{98} computed in this paper by a recurrence with gaps of 12 agrees with the value found from (3). Two errors in this paper should be noted. These occur in the recurrences for E_n with gaps of 6 and 12, where $2^{6\lambda}$ and $2^{12\lambda-1}$ should read $2^{6\lambda-2}$ and $2^{12\lambda}$, respectively.

[12] A complete table of G_n (n = 1 to 220), together with material by means of which the present table may be readily extended, is deposited in the library of the American Mathematical Society.

Table of Numerators of Bernoulli Numbers

N_{91}	N_{92}	N_{93}	N_{94}	N_{95}	N_{96}	N_{97}	N_{98}	N_{99}	N_{100}
								88527	498384
					346465	2	62753135	914861348	049428333
			14	5411555	752997582	269186825	110461193	004968400	414764928
4	857321	22258646	158277750	842544259	699690191	161532962	672553106	581010530	632140399
277269279	333523056	098436968	623758793	796131885	405750952	833665086	699893713	565220544	662108495
349192541	180131194	050639602	309386870	546196787	366871923	968359967	603153054	526400339	887457206
137304400	437347933	221816385	401397333	277987837	192340955	389321429	153311895	548429439	674968055
628629348	216431403	181596567	112823632	486638756	593486485	297588337	305590639	843908721	822617263
327468135	305730705	918515338	717478051	184149141	715370392	232986752	107017824	196349579	669621523
828402291	359015465	169946670	426522029	588783989	154894102	409765414	640241378	494069282	687568865
661683018	649285681	500599612	712001260	774511509	000406980	223476696	480484625	285662653	802302210
622451659	432317514	225742487	747920789	608733429	162521728	863199759	554578576	465989920	999132601
989595510	010686029	595012775	473711562	067517383	492501917	981611817	142115835	237253162	412697613
712915810	079324479	838387331	165031101	750706299	598012711	660735753	788960865	555666526	279391058
436238721	659634642	550474751	665618225	486822702	402163530	831323900	534532214	385826449	654527145
139546963	384809061	212260636	654329210	171672522	166516991	456495253	560982925	862863083	340515840
558655260	711319481	163500086	473605281	203106730	115122131	961837175	549798683	834096823	099290478
384328988	020030715	787417640	619696918	993581242	398542029	924312108	762705231	053048072	026350382
773219688	989009140	903770807	061316240	777825864	056286959	872915089	316611716	002986184	802884371
091443529	595170556	353228157	634857984	203487238	857727373	534970310	668749347	254693991	712359337
626531335	956196762	478339547	019071572	429479957	568402417	604331636	221458005	336699593	984274122
687951612	318625529	041472679	591940586	280273093	020319761	484174526	671217067	468906111	861159800
545946030	645723516	880890292	875558943	904025319	912636411	399721365	357943416	158296442	280019110
357929306	532076273	167353534	580878119	950569633	646719477	966337809	524984438	729034119	197888555
651006711	012244047	100797481	388321001	979493395	318166587	334021247	771831113	206322233	893671151

N_{101}	N_{102}	N_{103}	N_{104}	N_{105}	N_{106}	N_{107}	N_{108}	N_{109}	N_{110}
							11993122	5	8717064
				4708181	1856	4005	770108617	646413644	809960074
	110	2	12407390	368529492	110669947	748930070	858536443	023523531	651332043
2250	636644250	525292668	668433023	614110644	388268389	152861935	322964878	472659729	679796544
525326187	856903590	891404920	412711473	197951837	361040689	826766476	003618156	552574911	474420053
264545900	976481422	279427026	483696990	317202610	764027464	856180706	069559794	763686780	189621803
714460628	794879200	668969389	726334795	608341257	160460436	477227448	803117809	871700375	209941154
885135841	517231299	456388249	896412761	204206693	671923253	622268042	279608039	627426663	764877242
050444551	540994715	389889339	472587854	195241245	131176853	052745245	120818829	366507837	575579617
247116222	372334521	455604316	072142800	204360822	224087741	798242539	088000103	122353997	540185166
631411681	128669716	691573384	403373577	875910613	924378432	770546339	355036592	075891736	306094852
549780530	264196333	284678293	087021298	010433572	403442710	789899546	864877954	706811337	502223738
233516069	811025709	620100066	541061094	133227831	398247642	160341590	563564831	698248660	126111951
957534394	747746193	924361693	633377354	741097261	246902212	069109467	932363414	838754243	612525627
574922579	210786820	666444722	326966623	618833206	818749685	023517085	886811786	486778190	910517501
290608180	114369025	338743839	278849423	537519198	974336641	578618986	054601318	331522785	081576202
427520318	849897345	198221347	631924808	885812254	529240178	055969187	517206937	903661556	264770178
235621123	722531098	931651916	044397822	347219150	398124235	202731878	549605059	458651175	546608710
686109474	042760530	807651198	651135905	482005543	555437625	271685432	298307895	061469825	937474005
343887857	922656878	800935942	640812063	422997225	251481044	460708841	591515771	204821206	742969950
944611842	891556664	493038194	181221280	440204041	526024910	118506310	731031691	611273990	212404955
438698399	782168465	104759967	972334965	473518187	356131819	943865592	422489377	250663647	732473198
885295153	095563132	208073711	193338438	636442241	016670047	568791360	098686236	381261360	451623024
935574958	092311332	284671045	214107578	332621804	949661636	294244451	263367916	659950519	108934373
275021715	073097630	255047521	486417026	718967775	539964662	765746911	404512751	735730925	488641161
116120056	676251482	429204396	806166184	203938403	370375622	808994482	010916862	202117606	751158901
995036417	491663634	148980705	210160001	965710395	630863327	063783730	894129855	150672170	712323446
537079471	626858373	984836743	817890901	632762155	168696307	693703607	138281961	127523599	768306053

The corresponding denominators are as follows.

n	D_n	n	D_n
91	6	101	6
92	1410	102	281190
93	42	103	6
94	6	104	27030
95	12606	105	9225988926
96	868841610	106	3210
97	6	107	6
98	171390	108	15270994830
99	244713882	109	6
100	1366530	110	7590

Lehigh University.

BOOLEAN FUNCTIONS AND POINTS

By J. C. C. McKinsey

One of the important topics of algebraic geometry is that of the relations between points and rational curves. How many arbitrarily selected points, for example, can lie on a curve of degree n, and how many points are necessary to determine a curve of degree n? In this paper, I treat analogous problems in Boolean algebra. I take, for the analogue of the rational function in ordinary algebra, the Boolean function in Boolean algebra; that is to say, a function of Boolean variables which can be expressed by a finite number of applications of the Boolean operations $+$, $\times$, and $'$. For the analogue of the points of ordinary algebra, I define Boolean "points" as follows:

Definition 1. By an *n-space Boolean point* is meant an ordered set

$$(x_{1,1}, \cdots, x_{n,1}; z_1)$$

of Boolean elements. The ordered set $(x_{1,1}, \cdots, x_{n,1})$ is called the *abscissa* of the point. By saying that the point *satisfies* the function $f(x_1, \cdots, x_n)$, I mean that $f(x_{1,1}, \cdots, x_{n,1}) = z_1$; I also say that the function *passes through* the point, or that the point *lies on* the function, and use other terminology as in algebraic geometry.

My results are now developed in a series of theorems; a summary will be found at the end of the paper.[1]

Lemma to Theorem I. *If in the complete expansion of* 1 *in terms of n variables*

$$(1)\qquad x_1 \cdots x_n + x_1 \cdots x_{n-1}x_n' + \cdots + x_1' \cdots x_n'$$

we substitute u_iv_i for x_i and $u_i'v_i'$ for x_i', the resulting expression, namely,

$$u_1v_1 \cdots u_nv_n + u_1v_1 \cdots u_{n-1}v_{n-1}u_n'v_n' + \cdots + u_1'v_1' \cdots u_n'v_n',$$

is equal to

$$(u_1\Delta v_1)(u_2\Delta v_2) \cdots (u_n\Delta v_n).$$

Proof. The expression (1) is equal to

$$(2)\qquad (x_1 + x_1')(x_2 + x_2') \cdots (x_n + x_n').$$

Received December 16, 1935.

[1] Besides the ordinary operations of Boolean algebra, I make use of the operations $\circ$ and Δ defined by $a \circ b = ab' + a'b$, and $a\Delta b = ab + a'b'$. These operations are associative and commutative and mutually dual. A systematic discussion of them will be found in B. A. Bernstein's paper *Postulates for Boolean algebra involving the operation of complete disjunction* to appear in the Annals of Mathematics. I take this opportunity to acknowledge the valuable suggestions made by Professor Bernstein in connection with the present paper.

It is clearly immaterial whether the proposed substitution be made in (1) or in (2). Making it in (2), we have

$$(u_1v_1 + u'_1v'_1)(u_2v_2 + u'_2v'_2) \cdots (u_nv_n + u'_nv'_n),$$

or

$$(u_1\Delta v_1)(u_2\Delta v_2) \cdots (u_n\Delta v_n).$$

THEOREM I. *If r points $\{(x_{1,i}, \cdots, x_{n,i}; z_i)\}$, $i = 1, \cdots, r$, be given, there exists a function passing through the r points if and only if*

$$\sum_{j=1}^{r}\sum_{i=1}^{r}(z_i \cdot z_j)(x_{1,i}\Delta x_{1,j})(x_{2,i}\Delta x_{2,j}) \cdots (x_{n,i}\Delta x_{n,j}) = 0.$$

The r points determine a unique function if and only if in addition

$$\prod_{i=1}^{r}(x_{1,i} + \cdots + x_{n,i}) + \cdots + \prod_{i=1}^{r}(x'_{1,i} + \cdots + x'_{n,i}) = 0.$$

Proof. There exists a function passing through the r points if and only if there exist elements $A_1, \cdots, A_m$ (where, for typographical reasons, I write $m = 2^n$) such that

$$A_1x_{1,i} \cdots x_{n,i} + \cdots + A_mx'_{1,i} \cdots x'_{n,i} = z_i \qquad (i = 1, \cdots, r). \tag{1}$$

Conditions (1) are equivalent to

$$\begin{aligned} A'_1(x_{1,i} \cdots x_{n,i}z_i) + A_1(x_{1,i} \cdots x_{n,i}z'_i) + \cdots + A'_m(x'_{1,i} \cdots x'_{n,i}z_i) \\ + A_m(x'_{1,i} \cdots x'_{n,i}z'_i) = 0 \qquad (i = 1, \cdots, r). \end{aligned} \tag{2}$$

Conditions (2) are in turn equivalent to the single condition

$$\begin{aligned} A'\sum_{i=1}^{r}(x_{1,i} \cdots x_{n,i}z_i) + A_1\sum_{i=1}^{r}(x_{1,i} \cdots x_{n,i}z'_i) + \cdots \\ + A'_m\sum_{i=1}^{r}(x'_{1,i} \cdots x'_{n,i}z_i) + A_m\sum_{i=1}^{r}(x'_{1,i} \cdots x'_{n,i}z'_i) = 0. \end{aligned} \tag{3}$$

Thus a necessary and sufficient condition that there exist a function passing through the r points is that there exist a solution of (3) considered as an equation in $A_1, \cdots, A_m$. But there exists such a solution if and only if the following conditions obtain:

$$\begin{aligned} \sum_{i=1}^{r}(x_{1,i} \cdots x_{n,i}z_i) \cdot \sum_{i=1}^{r}(x_{1,i} \cdots x_{n,i}z'_i) &= 0, \\ \cdots \qquad \cdots \qquad &\cdots, \\ \sum_{i=1}^{r}(x'_{1,i} \cdots x'_{n,i}z_i) \cdot \sum_{i=1}^{r}(x'_{1,i} \cdots x'_{n,i}z'_i) &= 0. \end{aligned} \tag{4}$$

Conditions (4) are equivalent to the conditions

$$
(5) \qquad
\begin{aligned}
&\sum_{j=1}^{r}\sum_{i=1}^{r}(z_i\cdot z_j)[x_{1,i}x_{1,j}\cdots x_{n,i}x_{n,j}] = 0,\\
&\qquad\cdots\qquad\cdots\qquad\cdots,\\
&\sum_{j=1}^{r}\sum_{i=1}^{r}(z_i\cdot z_j)[x'_{1,i}x'_{1,j}\cdots x'_{n,i}x'_{n,j}] = 0.
\end{aligned}
$$

Conditions (5) are equivalent to the single condition

$$
(6) \qquad \sum_{j=1}^{r}\sum_{i=1}^{r}(z_i\cdot z_j)[x_{1,i}x_{1,j}\cdots x_{n,i}x_{n,j} + \cdots + x'_{1,i}x'_{1,j}\cdots x'_{n,i}x'_{n,j}] = 0.
$$

By the lemma, condition (6) is equivalent to

$$
(7) \qquad \sum_{j=1}^{r}\sum_{i=1}^{r}(z_i\cdot z_j)(x_{1,i}\Delta x_{1,j})\cdots(x_{n,i}\Delta x_{n,j}) = 0.
$$

Condition (7) is the condition specified in the first part of the theorem. For the second part of the theorem, we see[2] from (3) that we must have

$$
(8) \qquad
\begin{aligned}
&\left(\sum_{i=1}^{r}(x_{1,i}\cdots x_{n,i}z_i)\right)\Delta\left(\sum_{i=1}^{r}(x_{1,i}\cdots x_{n,i}z'_i)\right) = 0,\\
&\qquad\cdots\qquad\cdots\qquad\cdots,\\
&\left(\sum_{i=1}^{r}(x'_{1,i}\cdots x'_{n,i}z_i)\right)\Delta\left(\sum_{i=1}^{r}(x'_{1,i}\cdots x'_{n,i}z'_i)\right) = 0.
\end{aligned}
$$

To say that (8) holds is equivalent to saying that (4), and hence (7), holds, while in addition,

$$
(9) \qquad
\begin{aligned}
&\left(\sum_{i=1}^{r}(x_{1,i}\cdots x_{n,i}z_i)\right)'\cdot\left(\sum_{i=1}^{r}(x_{1,i}\cdots x_{n,i}z'_i)\right)' = 0,\\
&\qquad\cdots\qquad\cdots\qquad\cdots,\\
&\left(\sum_{i=1}^{r}(x'_{1,i}\cdots x'_{n,i}z_i)\right)'\cdot\left(\sum_{i=1}^{r}(x'_{1,i}\cdots x'_{n,i}z'_i)\right)' = 0.
\end{aligned}
$$

Taking the negatives of both sides of (9), we have

$$
(10) \qquad
\begin{aligned}
&\sum_{i=1}^{r}(x_{1,i}\cdots x_{n,i}z_i) + \sum_{i=1}^{r}(x_{1,i}\cdots x_{n,i}z'_i) = 1,\\
&\qquad\cdots\qquad\cdots\qquad\cdots,\\
&\sum_{i=1}^{r}(x'_{1,i}\cdots x'_{n,i}z_i) + \sum_{i=1}^{r}(x'_{1,i}\cdots x'_{n,i}z'_i) = 1.
\end{aligned}
$$

[2] A condition that a Boolean equation have a unique solution is given by A. N. Whitehead, *Memoir on the algebra of symbolic logic*, American Journal of Mathematics, vol. 23 (1901), pp. 140–150. A simpler derivation for this condition is given by B. A. Bernstein, *Note on the condition that a Boolean equation have a unique solution*, American Journal of Mathematics, vol. 54 (1932), pp. 417–418.

This is again equivalent to

$$(11)\qquad \sum_{i=1}^{r} x_{1,i} \cdots x_{n,i} = 1, \cdots, \sum_{i=1}^{r} x'_{1,i} \cdots x'_{n,i} = 1.$$

Taking the negatives of both sides of (11), we have

$$(12)\qquad \prod_{i=1}^{r} (x'_{1,i} + \cdots + x'_{n,i}) = 0, \cdots, \prod_{i=1}^{r} (x_{1,i} + \cdots + x_{n,i}) = 0.$$

Writing (12) as a single condition gives the condition specified in the second part of the theorem.

THEOREM II. *If P_1 is an arbitrary point, there exists a function passing through P_1. But if $r > 1$, there does not necessarily exist a function passing through r arbitrarily selected points.*

Proof. If we take $r = 1$ in Theorem I, then

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i}\Delta x_{1,j}) \cdots (x_{n,i}\Delta x_{n,j})$$
$$= (z_1 \cdot z_1)(x_{1,1}\Delta x_{1,1}) \cdots (x_{n,1}\Delta x_{n,1}) = 0.$$

Hence there always exists a function passing through one point. Indeed, such a function is $f(x_1, \cdots, x_n) = z_1$.

To show the second part of the theorem, we need to show that if $r > 1$, we can choose r points so that the expression

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i}\,\Delta x_{1,j}) \cdots (x_{n,i}\,\Delta x_{n,j})$$

does not vanish. But this can always be done by taking

$$z_1 = z'_2, \qquad x_{1,1} = x_{1,2}, \quad \cdots, \quad x_{n,1} = x_{n,2},$$

for then

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i}\,\Delta x_{1,j}) \cdots (x_{n,i}\,\Delta x_{n,j}) > (z_1 \cdot z_2)(x_{1,1}\,\Delta x_{1,2}) \cdots (x_{n,1}\,\Delta x_{n,2})$$
$$= (z_1 \cdot z'_1)(x_{1,1}\,\Delta x_{1,1}) \cdots (x_{n,1}\,\Delta x_{n,1}) = 1,$$

so that

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i}\,\Delta x_{1,j}) \cdots (x_{n,i}\,\Delta x_{n,j}) = 1 \neq 0.$$

It will be observed that in the last part of the proof just given we have employed two points $(x_{1,1}, \cdots, x_{n,1}; z_1)$, $(x_{1,1}, \cdots, x_{n,1}; z'_1)$, which are such that if a function passed through both of them it would not be single-valued. This is impossible for the functions we are considering. In this connection the following theorem is of interest.

THEOREM III. *Let $P_1, P_2, \cdots, P_r$ be a set of $r > 1$ arbitrarily selected points having distinct abscissas; then*

1. *In a two-element algebra, there always exists a function passing through* $P_1, \cdots, P_r$.

2. *In any other algebra, there does not necessarily exist a function passing through* $P_1, \cdots, P_r$.

Proof. 1. Consider, for a two-element algebra, the expression

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i} \Delta x_{1,j}) \cdots (x_{n,i} \Delta x_{n,j}).$$

Any term of this sum for which $i = j$ vanishes because $z_i \cdot z_i = 0$. Now let

$$T = (z_i \cdot z_j)(x_{1,i} \Delta x_{1,j}) \cdots (x_{n,i} \Delta x_{n,j})$$

be a term for which $i \neq j$. Since, by hypothesis, P_i and P_j have different abscissas, there must be some k such that $x_{k,i} \neq x_{k,j}$. But $x_{k,i}$ and $x_{k,j}$ are each either 0 or 1, hence $x_{k,i} \Delta x_{k,j} = 0$, since $0\Delta 1 = 1\Delta 0 = 0$. Hence $T = 0$. Thus we have

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i} \Delta x_{1,j}) \cdots (x_{n,i} \Delta x_{n,j}) = 0.$$

Hence, by Theorem I, there exists a function passing through $P_1, \cdots, P_r$.

2. An algebra which is not a two-element algebra must contain an element a which is different from 1 and 0. Consider now $P_1, P_2, \cdots, P_r$, where $P_1 = (1, 1, \cdots, 1; 1)$, $P_2 = (a, a, \cdots, a; 0)$ and $P_3, \cdots, P_r$ are arbitrarily chosen. Then

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i} \Delta x_{1,j}) \cdots (x_{n,i} \Delta x_{n,j}) > (z_1 \cdot z_2)(x_{1,1} \Delta x_{1,2}) \cdots (x_{n,1} \Delta x_{n,2})$$
$$= (1 \cdot 0)(1 \Delta a) \cdots (1 \Delta a) = (1)(a) \cdots (a) = a \neq 0.$$

Hence

$$\sum_{j=1}^{r} \sum_{i=1}^{r} (z_i \cdot z_j)(x_{1,i} \Delta x_{1,j}) \cdots (x_{n,i} \Delta x_{n,j}) \neq 0,$$

and thus by Theorem I there does not exist a function passing through the r points.

The first part of Theorem III is of some interest in connection with the logic of propositions, since the logic of propositions is a two-element Boolean algebra. The theorem then asserts that we can find a function of the r propositional variables $p_1, \cdots, p_r$ which will assume any pre-assigned sets of truth-values for any pre-assigned sets of truth-values of $p_1, \cdots, p_r$; this property has been called "symbolic completeness" by C. I. Lewis.[3] The theorem is proved in another way by E. L. Post.[4]

[3] Lewis and Langford, *Symbolic Logic*, p. 231.

[4] *Introduction to a general theory of propositions*, Amer. Jour. Math., vol. 43 (1921), pp. 163–185.

The following theorem sets a lower limit to the number of points that can determine a unique function.

THEOREM IV. *If $r < 2^n$, then there exists no set of r points*

$$\{(x_{1,i}, \cdots, x_{n,i}; z_i)\}, \qquad i = 1, \cdots, r,$$

which determines a unique function.

If $r \geqq 2^n$, there exist sets of r points which determine unique functions as well as sets of r points which do not determine unique functions.

Proof. To show that $r < 2^n$ points never determine a unique function, it is sufficient, by Theorem I, to show that the expression

$$\prod_{i=1}^{r} (x_{1,i} + \cdots + x_{n,i}) + \cdots + \prod_{i=1}^{r} (x'_{1,i} + \cdots + x'_{n,i}) \tag{1}$$

does not vanish for any choice of the x's if $r < 2^n$.

Let us find the discriminants in the complete expansion of (1), putting $t = 2^{nr}$,

$$A_1 x_{1,1} x_{1,2} \cdots x_{n,r} + \cdots + A_t x'_{1,1} x'_{1,2} \cdots x'_{n,r}. \tag{2}$$

These A's are all either 0 or 1, for they are found by substituting 0 and 1 for the x's in (1). Suppose now that some $A_k = 0$. This means that there exists a set of values (all 0 or 1) for the x's in (1) which makes (1) vanish. Let us denote such a set of values of the x's by y's. Then we have

$$\prod_{i=1}^{r} (y_{1,i} + \cdots + y_{n,i}) + \cdots + \prod_{i=1}^{r} (y'_{1,i} + \cdots + y'_{n,i}) = 0. \tag{3}$$

Hence, in particular,

$$\prod_{i=1}^{r} (y_{1,i} + \cdots + y_{n,i}) = 0. \tag{4}$$

Each factor of the product (4), however, is either 0 or 1. Hence there must be at least one factor which vanishes. Suppose this is the first factor; then we have

$$y_{1,1} + \cdots + y_{n,1} = 0. \tag{5}$$

Hence

$$y_{1,1} = \cdots = y_{n,1} = 0, \tag{6}$$

and

$$y_{1,1} + \cdots + y_{n-1,1} + y'_{n,1} = 1, \cdots, y'_{1,1} + \cdots + y'_{n-1,1} + y'_{n,1} = 1. \tag{7}$$

Substituting (5) and (7) in (3), we obtain

$$\prod_{i=2}^{r} (y_{1,i} + \cdots + y_{n-1,i} + y'_{n,i}) + \cdots + \prod_{i=2}^{r} (y'_{1,i} + \cdots + y'_{n,i}) = 0. \tag{8}$$

We can now repeat the argument using the next product

$$\prod_{i=2}^{r} (y_{1,i} + \cdots + y_{n-1,i} + y'_{n,i}) = 0, \tag{9}$$

obtaining, say,

$$\prod_{i=3}^{r} (y_{1,i} + \cdots + y_{n-2,i} + y'_{n-1,i} + y_{n,i}) + \cdots + \prod_{i=3}^{r} (y'_{1,i} + \cdots + y'_{n,i}) = 0. \tag{10}$$

At each step we reduce by one the number of terms in the sum (of products) represented in (1); at each step we assign values to n of the y's. Hence at the end of r steps we shall have assigned values to all of the y's. But r steps will not have exhausted the terms of (1); for (1) has 2^n terms, and by hypothesis $r < 2^n$. Hence, at the end of the r-th step, we shall be led to the contradiction $1 = 0$. Therefore it is not the case that there exists a k so that $A_k = 0$. But we have said that $A_k = 0$ or $A_k = 1$. Hence, for all k, $A_k = 1$. Then (2) becomes

$$x_{1,1}x_{1,2} \cdots x_{n,r} + \cdots + x'_{1,1}x'_{1,2} \cdots x'_{n,r} \tag{11}$$

which equals 1. Hence, from (1),

$$\prod_{i=1}^{r} (x_{1,i} + \cdots + x_{n,i}) + \cdots + \prod_{i=1}^{r} (x'_{1,i} + \cdots + x'_{n,i}) = 1 \neq 0 \tag{12}$$

for $r < 2^n$. Thus $r < 2^n$ points never determine a unique function.

To show that it is always possible to find 2^n points that determine a unique function, it is sufficient to consider the 2^n points

$$(1, \cdots, 1; 1), \cdots, (0, \cdots, 0; 1). \tag{13}$$

These points clearly satisfy the first condition of Theorem I, since $z_i \cdot z_j = 1 \cdot 1 = 0$. Also, each product in the second condition of Theorem I will contain a vanishing factor, and hence will itself vanish. Thus the second condition is satisfied.

For $r > 2^n$, we can choose 2^n of the points as those specified in (13) and the others as $(u, v, \cdots, w; 1)$, where $u, v, \cdots, w$ are arbitrary.

To show that for $r \geqq 2^n$ we can always find r points which do *not* determine a unique function, it is sufficient to take r points which do not satisfy the first condition of Theorem I. Hence the points can be taken as in the proof of Theorem II.

Summary. Given a set of r n-space Boolean points $\{(x_{1,i}, \cdots, x_{n,i}; z_i)\}$; if $r = 1$, a function can always be found passing through the point. If $r > 1$, a function can sometimes, but not always, be found passing through the points of the set; with the important exception that, if the algebra is a two-element algebra, and the r points have distinct abscissas, such a function can always be found.

If $r < 2^n$, the set of points will never determine a unique function. If $r \geqq 2^n$, the set of points will sometimes, but not always, determine a unique function.

THE UNIVERSTY OF CALIFORNIA.

THE NULL DIVISORS OF LINEAR RECURRING SERIES

By Morgan Ward

1. Let

$$(u) \qquad u_0, u_1, u_2, \cdots, u_n, \cdots$$

be a particular solution of the difference equation

$$(1.1) \qquad \Omega_{n+k} = c_1\Omega_{n+k-1} + \cdots + c_k\Omega_n,$$

where $u_0, \cdots, u_{k-1}, c_1, \cdots, c_k$ are given rational integers and $c_k \neq 0$. If all of the terms of (u) beyond a certain point are divisible by a given integer m, then m will be said to be a *null divisor* of (u), and (u) a *null sequence* modulo m. In this case there is an integer ν called the *numeric* of (u) modulo m such that[1]

$$u_n \equiv 0 \pmod{m}, \qquad n \geqq \nu, \qquad u_{\nu-1} \not\equiv 0 \pmod{m}.$$

In a previous paper,[2] I have solved the problem of determining the numeric of (u) given its k initial values, the recurrence (1.1) and the null divisor m. In this paper I propose to determine all of the null divisors of (u).[3]

If a and b are null divisors, then ab is also a null divisor provided a and b are co-prime. It suffices then to consider only the case when m is a power of a prime. If p is a prime null divisor of (u), the exponent of the highest power of p dividing all terms of (u) with large suffixes will be called the index of p in (u). If, for example, from a certain point on all terms of (u) are divisible by p^2 but not by p^3, p is of index two.

2. My main results are summarized in the following two theorems.

Theorem 1. *If in the difference equation* (1.1) *we have*

$$(2.1) \qquad c_k \equiv c_{k-1} \equiv \cdots \equiv c_{k-s+1} \equiv 0 \pmod{p}, \qquad c_{k-s} \not\equiv 0 \pmod{p},$$

where p is a prime, and if d_n denotes the greatest common divisor of the $k - s$ consecutive terms

$$u_{n+s}, u_{n+s+1}, \cdots, u_{n+k-1} \qquad (n \geqq 0)$$

Received March 10, 1936.

[1] We exclude from consideration the trivial common divisors of $u_0, u_1, \cdots, u_{k-1}$.

[2] *The arithmetical theory of linear recurring series*, Trans. Am. Math. Soc., vol. 35 (1933), pp. 600–628. This paper will be cited here as Theory.

[3] The present paper is a condensation and completion of some earlier results on the same subject. Cf. Abstract 22, Bulletin Am. Math. Soc., vol. 41 (1935), p. 24. Very recently, Abstract 11, Bulletin Am. Math. Soc., vol. 42 (1936), p. 25, Mr. Marshall Hall has given some results on null divisors which we shall discuss in the course of the paper.

of (u), *then a necessary and sufficient condition that the index of* p *in* (u) *be* λ *is that* p^λ *be the highest power of* p *dividing* $d_{\lambda s}$. *If* $d_{\lambda s} \equiv 0 \pmod{p^{\lambda+1}}$, *then* $u_n \equiv 0 \pmod{p^{\lambda+1}}$, $n \geqq (\lambda + 1)s$.

It follows from this theorem that the prime null divisors of (u) are common divisors of c_k and u_{k-1}, and that if λ is the index of p in (u), then the numeric of (u) modulo p^λ is less than or equal to λs.

Let $\Delta(u)$ denote the k-rowed persymmetric determinant[4]

$$\Delta(u) = \begin{vmatrix} u_0, & u_1, & \cdots, & u_{k-1} \\ u_1, & u_2, & \cdots, & u_k \\ \cdots & \cdots & \cdots & \cdots \\ u_{k-1}, & u_k, & \cdots, & u_{2k-2} \end{vmatrix}. \tag{2.2}$$

THEOREM 2. *If* $\Delta(u)$ *is not equal to zero, and if the coefficients* $c_1, c_2, \cdots, c_k$ *of* (1.1) *are co-prime, then for any prime null divisor* p *of* (u), *the index of* p *in* (u) *is less than or equal to the highest power of* p *contained in the first elementary divisor of the matrix associated with the determinant* $\Delta(u)$.

If $\Delta(u)$ vanishes, then (u) satisfies a difference equation of order[5] less than k. If $c_1, c_2, \cdots, c_k$ have p as a common factor, there is no limit to the size of the index[6] of p.

These two theorems give a simple and practicable way of determining both all the null divisors of any given linear recurring series, and the numeric of any null divisor.[7] For if ν_r is the numeric of (u) modulo p^r, and if p^σ is the power of p associated with $\Delta(u)$ described in Theorem 2, we have the inequalities $\nu_r \leqq rs \leqq \sigma s$.

3. My proof of Theorem 1 will be based on the following

LEMMA. *Under the hypotheses of Theorem* 1, *if* m_0 *is any positive integer, then* $d_{m_0} \equiv 0 \pmod{p}$ *when and only when* $d_0 \equiv 0 \pmod{p}$.

For let A, B, C, D and E denote rational integers, and m any modulus. Clearly

(i) If $A \equiv B \pmod{m}$, then[8] $(C, A) \equiv 0 \pmod{m}$ when and only when

$$(C, B) \equiv 0 \pmod{m}.$$

[4] For other arithmetical properties of $\Delta(u)$, see Theory. On pp. 604, 608 of Theory the element in the lower right corner of $\Delta(u)$ should be u_{2k-2} instead of u_{2k-1}.

[5] See §5 of this paper.

[6] This result is given by Mr. Marshall Hall, abstract cited. See also §5 of this paper.

[7] The method for determining the numeric which I have given in Theory, pp. 614–616 is only of theoretical interest.

[8] We use the customary notation $(x, y, z, \cdots)$ for the greatest common divisor of $x, y, z, \cdots$.

(ii) If $(D, m) = 1$, then $(E, C) \equiv 0 \pmod{m}$ when and only when

$$(DE, C) \equiv 0 \pmod{m}.$$

Now take p for the modulus m, and let

$$A = u_{n+k}, \qquad B = c_1 u_{n+k-1} + c_2 u_{n+k-2} + \cdots + c_{k-s} u_{n+s},$$

$$C = (u_{n+s+1}, u_{n+s+2}, \cdots, u_{n+k-1}), \qquad D = c_{k-s}, \qquad E = u_{n+s}.$$

Then $(A, C) = d_{n+1}$, $(B, C) = (DE, C)$ and $(E, C) = d_{n+1}$. Also by (2.1), $A \equiv B \pmod{p}$ and $(D, p) = 1$. Therefore, by (i) and (ii), $d_{n+1} \equiv 0 \pmod{p}$ when and only when $d_n \equiv 0 \pmod{p}$, and the lemma is evident.

4. **Proof of Theorem** 1. First, suppose that the index λ of p in (u) is zero. Then by the lemma, $(p, d_0) = 1$, and conversely if $(p, d_0) = 1$, the index is zero. By the lemma again, if[9] $p \mid d_0$, then $p \mid u_n$, $n = s, s + 1, \cdots$. Therefore the theorem is true when $\lambda = 0$.

Assume that the theorem is true when $\lambda = k$. Then it is also true when $\lambda = k + 1$. For by our assumption and the theorem itself, $\lambda > k$ when and only when $u_n \equiv 0 \pmod{p^{k+1}}$, $n \geqq (k + 1)s$. Write

$$u_{n+(k+1)s} = p^{k+1} u'_n \qquad (n = 0, 1, 2, \cdots). \tag{4.1}$$

Then (u') is a solution of (1.1).

Let $d'_n = (u'_{n+s}, u'_{n+s+1}, \cdots, u'_{n+k-1})$. By the lemma again, the index of p in (u') is zero when and only when $(p, d'_0) = 1$ and if $p \mid d'_0$, then $p \mid u'_n$, $n \geqq s$.

Therefore by (4.1), the index of p in (u) is $k + 1$ when and only when $p^{k+1} \parallel d_{(k+1)s}$, and if $p^{k+2} \mid d_{(k+1)s}$, then $u_n \equiv 0 \pmod{p^{k+2}}$, $n \geqq (k + 2)s$. Thus Theorem 1 follows by induction.

5. We preface our proof of Theorem 2 by some results from the algebraic theory of recurring series which are of arithmetical importance.[10]

Let

$$F(x) = x^k - c_1 x^{k-1} - c_2 x^{k-2} - \cdots - c_k$$

be the polynomial associated with the recurrence (1.1). Then if

$$F_r(x) = x^r - c_1 x^{r-1} - \cdots - c_r \qquad (r = 0, 1, \cdots, k),$$

so that $F_0(x) = 1$, $F_k(x) = F(x)$, I have called the polynomial

$$U(x) = u_0 F_{k-1}(x) + u_1 F_{k-2}(x) + \cdots + u_{k-1} F_0(x) \tag{5.1}$$

[9] We write $p \mid x$ for p divides x. If the highest power of p dividing x is p^ρ, we shall write $p^\rho \parallel x$.

[10] Most of these results may be found in chapter XII, §II of the well known *Treatise on the Theory of Determinants*, by Muir and Metzler, Albany, 1930.

the generator of the sequence (u). (Theory, p. 606.) We have in fact for sufficiently large values of x the identity

$$U(x)/F(x) = \sum_{n=0}^{\infty} u_n/x^{n+1}. \tag{5.2}$$

Furthermore,[11]

$$\Delta(u) = (-1)^{k(k-1)/2}\ \mathrm{Res}\ \{U(x), F(x)\}. \tag{5.3}$$

It is obvious from (5.2) and (5.3) that if $\Delta(u)$ vanishes, (u) satisfies a difference equation of order less than k.

If we write the left side of (5.2) as

$$x^{-k}U(x)(1 - c_1/x - \cdots - c_k/x^k)^{-1},$$

it is also obvious that if $c_1, \cdots, c_k$ have a common divisor m, all terms of (u) beyond a certain point are divisible by m^λ for any preassigned value of λ. Thus[12] the index in (u) of a prime dividing $c_1, \cdots, c_k$ is unbounded.

Finally, since by (5.1)

$$U(x) = u_0x^{k-1} + (u_1 - c_1u_0)x^{k-2} + \cdots + (u_{k-1} - c_1u_{k-2} - \cdots - c_{k-1}u_0),$$

the coefficients of $U(x)$ are relatively prime when and only when

$$u_0, u_1, \cdots, u_{k-1}$$

are relatively prime.

6. We conclude with the proof of Theorem 2. I have shown in Theory that the numeric of (u) modulo p^λ is the least value of N such that

$$x^N U(x) \equiv 0 \qquad (\mathrm{modd}\ p^\lambda, F(x)). \tag{6.1}$$

But I have also shown[13] that for fixed polynomials $U(x)$ and $F(x)$ (where $F(x)$ is primary and the coefficients of $U(x)$ are not divisible by p) the congruence

$$Y(x)U(x) \equiv 0 \qquad (\mathrm{modd}\ p^\lambda, F(x)) \tag{6.2}$$

has solutions $Y(x)$ not divisible by p only if $\lambda \leqq \sigma$, where p^σ is the highest power of p dividing the first elementary divisor $\bar{e}$ of the matrix of the resultant of $U(x)$ and $F(x)$. But it is easily shown that $\bar{e}$ is also the first elementary divisor of the matrix of $\Delta(u)$, so that the theorem is established.

[11] A proof of this formula is indicated in Theory, pp. 608–609. The sign, however, is incorrectly given there as $(-1)^k$. The result goes back to Netto, Journal für Math., vol. 106 (1895), pp. 33–49; Muir's *History*, vol. III, p. 326.

[12] Marshall Hall, abstract cited.

[13] Trans. Am. Math. Soc., vol. 35 (1933), pp. 254–260.

7. In the congruence (6.1), let us regard p, N and $U(x)$ as unknown, and λ and $F(x)$ as pre-assigned. On observing that Res $\{x^N, F(x)\} = \pm c_k^N$, we see that if p divides c_k, then by the result just stated for the congruence (6.2) we can choose a polynomial $U(x)$ not divisible by p to satisfy (6.1) provided that N is taken so large that the first elementary divisor[14] of the matrix of Res $\{x^N, F(x)\}$ is divisible by p^λ. In the corresponding sequence (u), the index of p is $\geqq \lambda$. Hence no upper limit exists for the indices of p in *all* the solutions of (1.1) admitting p as a null divisor.[15]

California Institute of Technology.

[14] It is not difficult to show that the power of p which divides this elementary divisor increases with N.

[15] Marshall Hall, abstract cited.

THE SIMPLE GROUP OF ORDER 25920

By J. S. Frame

1. Among the 53 known simple groups of composite order less than one million,[1] 42 may be represented as linear fractional modular groups on two or three variables. There are in addition three other alternating groups, and three multiply transitive groups. Of the five remaining groups, three are hyperorthogonal groups on three variables, whose irreducible representations were discussed in a recent paper.[2] The other two, of orders 25920 and 979200 respectively, may be defined as the "abelian linear groups"[1] $A(4, 3)$ and $A(4, 4)$. The first of these is also isomorphic to the hyperorthogonal group $HO(4, 4)$ on four variables in a modular field of four marks. Thus it is the smallest example both of the "abelian linear group" and of the hyperorthogonal group on more than three variables. We are interested here in its properties from the latter point of view, and shall obtain the complete table of characters of its irreducible representations.

The group $HO(m, q^2)$ may be defined as the quotient group with respect to its central of the special unitary group G_m, consisting of matrices of degree m and determinant 1, with coefficients from a finite field $GF(q^2)$ of q^2 marks. Here $q = p^e$ is a power of the prime p, and "conjugate imaginaries" are defined by the equation $\bar{x} = x^q$. We may think of the transformations of the group G_m as operating on a set of vectors in an m-dimensional space where the coördinates are marks of the $GF(q^2)$. All multiples of a given vector will be said to form a *ray*. The inner product, $(a \mid b) \equiv \sum_{i=1}^{m} \bar{a}_i b_i$, of two vectors a and b is invariant under each transformation of the unitary group G_m, so that the *isotropic* vectors—those for which $(a \mid a) = 0$—are permuted among themselves, and so are the remaining *non-isotropic* vectors, for which $(a \mid a) \neq 0$. It has been shown[3] that the permutation groups thus induced on the rays of each of the two types are transitive. If a single vector be selected from each ray, these vectors undergo a *monomial substitution* under the group G_m, with multipliers which, although appearing as marks from the $GF(q^2)$, may be replaced by suitably chosen $(q^2 - 1)$-th roots of unity from the field of complex numbers. It has also been shown that for $m = 3$ the permutation group on the isotropic rays is doubly transitive, and the corresponding monomial groups either have just two irreducible components, or are irreducible. In this way more than

Received February 20, 1936.

[1] L. E. Dickson, *Linear Groups*, 1901, p. 309.

[2] J. S. Frame, *Some irreducible monomial representations of hyperorthogonal groups*, this Journal, vol. 1 (1935), p. 442.

[3] J. S. Frame, *Unitäre Matrizen in Galoisfeldern*, Commentarii Mathematici Helvetici, vol. 7 (1935), p. 97.

half of the irreducible representations of $HO(3, q^2)$ could be found. But the problem of reduction is more complicated when $m > 3$, since the permutation group on the isotropic rays is then only simply transitive, and has always just three irreducible components, while the corresponding monomial representations have either three or two irreducible components. This problem, as well as the problem of reducing the permutation group on the non-isotropic rays, can best be solved by studying the hermitian invariants of these groups.

2. We know that the number of irreducible components of a linear group is equal to the number of linearly independent hermitian invariants. If we let the variable x_a correspond to the vector a from the isotropic ray R_a, we find that the three hermitian forms $\sum_a \bar{x}_a x_a$, $\sum \bar{x}_a x_b$ with $(a \mid b) \neq 0$, and $\sum \bar{x}_a x_b$ with $(a \mid c) = 0$ are linearly independent, and are invariant under the permutation group P_m on the isotropic rays. In order to display the reducibility of P_m, these hermitian invariants must be transformed to diagonal form. Let us denote by I, J, and H respectively the matrices of their coefficients, and also introduce the matrix $K = I + J + H$, in which each element is 1. Each of these matrices is symmetric and has $n = Q_m Q_{m-1}/Q_2$ rows and columns, if we set $Q_m = q^m - (-1)^m$. I is the unit matrix, and H is a matrix having $h = q^2 Q_{m-2} Q_{m-3}/Q_2$ 1's in each row and column, with zeros elsewhere. The inner product of two rows of H which correspond to the variables x_a and x_b will be $h_0 = Q_2 + q^4 Q_{m-4} Q_{m-5}/Q_2$ if $(a \mid b) = 0$, and $h_1 = Q_{m-2} Q_{m-3}/Q_2$ if $(a \mid b) \neq 0$. Hence

$$H^2 = hI + h_1 J + h_0 H = (h - h_1)I + (h_0 - h_1)H + h_1 K.$$

Now since $HK = hK$, the matrix H satisfies the minimal equation

$$(H - hI)\,[H^2 - (h_0 - h_1)H - (h - h_1)I] = 0.$$

The roots of this equation are h, $(-q)^{m-3} - 1$, and $(-q)^{m-2} - 1$. Knowing that H has in all n roots, with sum zero, and with sum of squares equal to hn, the multiplicities of the three distinct roots may be calculated. They are found to be 1 for the root h, $q^3 Q_{m-1} Q_{m-2}/Q_2 Q_1$ for the root $(-q)^{m-3} - 1$, and $q^2 Q_m Q_{m-3}/Q_2 Q_1$ for the root $(-q)^{m-2} - 1$. It is found similarly that the matrix J satisfies the minimal equation

$$[J - (n - 1 - h)I]\,[J^2 - (h_1 - h_0 - 2)J - (h - h_0 - 1)I] = 0,$$

whose roots q^{2m-3}, $-(-q)^{m-3}$, and $-(-q)^{m-2}$ occur respectively with the same multiplicities as those of H. These multiplicities are precisely the degrees of the irreducible components P_m.

THEOREM 1. *The representation of the hyperorthogonal group $HO(m, q^2)$ as a permutation group P_m of degree $Q_m Q_{m-1}/Q_2$ on the isotropic GF-rays in m dimensions has just three irreducible components, whose degrees are* 1, $q^3 Q_{m-1} Q_{m-2}/Q_2 Q_1$ *and* $q^2 Q_m Q_{m-3}/Q_2 Q_1$ *respectively.* (*When $m = 3$, the last of these vanishes.*)

The $q^{m-1} Q_m/Q_1$ non-isotropic rays R_a, with $(a \mid a) \neq 0$, are permuted tran-

sitively among themselves by the transformations of G_m. The subgroup leaving one ray invariant permutes the others in q transitive sets; except that when $m = 3$ and 3 divides $q + 1$, there are $q + 2$ transitive sets. Hence the permutation group Φ_m on these rays has just $q + 1$ (or $q + 3$) irreducible components. The corresponding invariant hermitian forms are the unit form $\sum \bar{x}_a x_a$, and the q forms $\sum \bar{x}_a x_b$, where $(a \mid b)(b \mid a)/(a \mid a)(b \mid b) = k$, and k is in turn each one of the q marks from the $GF(q)$. (In the special case mentioned above, the form for which $k = 0$ splits into three separate parts.) Proceeding as for the permutation group P_m, we now have $q + 1$ different matrices instead of three to reduce to diagonal form. The computations are too complicated to be given here at length, so we shall merely state the results without proof.

Theorem 2. *The representation of the hyperorthogonal group $HO(m, q^2)$ as a permutation group Φ_m of degree $q^{m-1}Q_m/Q_1$ on the non-isotropic GF-rays in m dimensions has in general $q + 1$ distinct irreducible components. These include: the identity of degree 1 and one representation of degree $q^3Q_{m-1}Q_{m-2}/Q_2Q_1$ which are components of P_m, and in addition one of degree qQ_mQ_{m-1}/Q_2Q_1 if q is odd, $[q/2] - 1$ of degree Q_mQ_{m-1}/Q_2, and $[q/2]$ of degree Q_mQ_{m-1}/Q_1^2. When $m = 3$, and 3 divides $q + 1$, one of the last type splits into three components of equal degree.*

For the monomial substitutions on the isotropic rays the unit hermitian form is invariant, and also the form $\sum e_{ab}\bar{x}_a x_b$, with $(a \mid b) = 0$, $a \neq b$, and suitably chosen coefficients $e_{ab} = \bar{e}_{ba}$. There is a third bilinear form $\sum c_{ab}\bar{x}_a x_b$, $(a \mid b) = k \neq 0$, where c_{ab} is the complex $(q^2 - 1)$-th root of unity which is made to correspond to the mark $1/k$; but this is hermitian only when all the coefficients c_{ab} are $(q + 1)$-th roots of unity. In this case the degrees of the three irreducible components are Q_mQ_{m-3}/Q_2, $q^{m-2}Q_m/Q_1$, and $q^{m-3}Q_m/Q_1$. Otherwise, these last two components are to be replaced by a single one of degree $q^{m-3}Q_m$. In either case the first component vanishes when $m = 3$. The monomial substitutions on the non-isotropic rays have similar hermitian invariants, $q + 1$ or q in number, depending on which roots of unity are used to correspond to the marks of the finite field. Two component representations of degree $q^{m-2}Q_m/Q_1$ and $q^{m-3}Q_m/Q_1$ or a single one of degree $q^{m-3}Q_m$ are the same as above. The remaining $q - 1$ components in the reduction are of degree $q^{m-3}Q_m$ and $(q - 1)q^{m-3}Q_m/Q_1$, except for one of degree $q^{m-2}Q_m/Q_1$ if q is odd.

3. In order to illustrate how these reducible representations split up in general into irreducible components, we consider in this paper the simplest example of these groups, namely, the simple group $HO(4, 4)$ of order 25920. We denote the marks of the $GF(4)$ by the symbols $0, 1, \omega, \omega^2$, where

$$1 + \omega + \omega^2 \equiv 0 \pmod 2.$$

Then the 135 four-dimensional isotropic vectors (the null vector not included), of which either two components or none are zero, line up in 45 isotropic rays; and the remaining 120 vectors, of which one or three components are zero,

form 40 non-isotropic rays. Each set of rays is permuted transitively when transformed by the matrices of G_4. In each case the subgroup leaving one ray R_a fixed permutes the remaining rays *in two transitive sets*, so the permutation groups P_4 and Φ_4 each have *three* irreducible components. The corresponding monomial representations, two conjugate imaginary representations of degree 45, and two of degree 40, each have just two irreducible components.

The operations of the group G_4 fall into 20 sets of conjugates. In the following analysis we display the matrices of 10 of the sets in normal forms generalized for all m, and evaluate the indices of the sets for the particular group $HO(4, 4)$, for which $m = 4$ and $q = 2$. To obtain the remaining 10 conjugate sets, we examine the subgroups permutable with selected matrices. We denote by d the order of the central of G_m, and abbreviate $q^m - (-1)^m$ by the symbol Q_m.

(1) Only the d matrices of the central of G_m—those which correspond to the identity in the quotient group $H_m \equiv HO(m, q^2)$—can leave m linearly independent isotropic rays invariant, and they leave $Q_mQ_{m-1}/Q_2 = 45$ isotropic, and $q^{m-1}Q_m/Q_1 = 40$ non-isotropic rays absolutely invariant. These numbers are the respective degrees of the reducible monomial representations we are investigating.

(2) The $Q_mQ_{m-1}/Q_1 = 45$ matrices of order p, of the form $(\delta_{ij} + \epsilon\bar{a}_ia_j)$, where $(a \mid a) = \epsilon + \bar{\epsilon} = 0$, leave absolutely invariant the

$$1 + q^2Q_{m-2}Q_{m-3}/Q_2 = 13$$

isotropic, and $q^{m-1}Q_{m-2}/Q_1 = 8$ non-isotropic, rays orthogonal to the vector a.

(3) The $qQ_mQ_{m-1}Q_{m-2}Q_{m-3}/Q_2Q_1 = 270$ matrices of order p, of the form $(\delta_{ij} + \bar{a}_ib_j - \bar{b}_ia_j)$, where $(a \mid a) = (a \mid b) = (b \mid b) = 0$, $a \neq kb$, constitute a single set of conjugate matrices leaving absolutely invariant

$$1 + q^2 + q^4Q_{m-4}Q_{m-5}/Q_2 = 5$$

isotropic, and $q^{m-1}Q_{m-4}/Q_1 = 0$ non-isotropic rays.

(4) The $q^{m-2}Q_mQ_{m-1}Q_{m-2}/Q_1 = 540$ matrices of order p^2 (if $p = 2$), or p (if $p > 2$), of the form $(\delta_{ij} + \bar{b}_ia_j - \bar{a}_ib_j + \epsilon\bar{a}_ia_j)$, where $(a \mid a) = (b \mid a) = (b \mid b) + \epsilon + \bar{\epsilon} = 0$, constitute a single set of conjugate matrices when $(b \mid b) \neq 0$, $m > 3$, which leave absolutely invariant $1 + q^2Q_{m-3}Q_{m-4}/Q_2 = 1$ isotropic and $q^{m-2}Q_{m-3}/Q_1 = 4$ non-isotropic rays.

(5, 6) The $q^{m-1}Q_m/Q_1 = 40$ matrices $[\theta\delta_{ij} + \theta(\theta^{-m} - 1)\bar{c}_ic_j/(c \mid c)]$, where $(c \mid c) \neq 0$, $\theta\bar{\theta} = 1$, $\theta \neq 1$, form for each admissible θ a set of conjugate matrices whose periods divide Q_1. Each of these leaves relatively invariant with factor θ just $Q_{m-1}Q_{m-2}/Q_2 = 9$ isotropic rays, and $q^{m-2}Q_{m-1}/Q_1 = 12$ non-isotropic rays, and also with factor θ^{1-m} the non-isotropic ray R_c itself. There are $Q_1/d - 1 = 2$ such sets in H_m.

(7, 8) The $q^{m-1}Q_mQ_{m-1}Q_{m-2}/Q_1^2 = 360$ matrices

$$[\theta\delta_{ij} + \theta\epsilon\bar{a}_ia_j + \theta(\theta^{-m} - 1)\bar{c}_ic_j/(c \mid c)],$$

where $\epsilon + \bar{\epsilon} = (a \mid a) = (c \mid a) = 0$, $\epsilon a(c \mid c) \neq 0$, $\theta\bar{\theta} = 1$, $\theta \neq 1$, form for each admissible θ a set of conjugate matrices whose periods divide pQ_1, each

of which leaves relatively invariant with factor θ certain $1 + q^2Q_{m-3}Q_{m-4}/Q_2 = 1$ isotropic and $q^{m-2}Q_{m-3}/Q_1 = 4$ non-isotropic rays, and also with factor θ^{1-m} the non-isotropic ray R_c itself. There are $Q_1/d - 1 = 2$ such sets in H_m.

(9, 10) The $q^{2m-4}Q_mQ_{m-1}Q_{m-2}Q_{m-3}/Q_1^2 = 2160$ matrices

$$[\theta(\delta_{ij} + \bar{b}_ia_j - \bar{a}_ib_j + \epsilon a_ia_j) + \theta(\theta^{-m} - 1)\bar{c}_ic_j/(c \mid c)],$$

where $(a \mid a) = (a \mid b) = (a \mid c) = (b \mid c) = (b \mid b) + \epsilon + \bar{\epsilon} = 0$, $(c \mid c) \neq 0$, $\theta\bar{\theta} = 1$, $\theta \neq 1$, form for each admissible θ a set of conjugates whose periods divide p^2Q_1. Any such matrix can leave invariant only the one ray R_c, and relatively invariant only those rays such as R_a which are orthogonal to a, b, and c. Relatively invariant with factor θ are $1 + q^2Q_{m-4}Q_{m-5}/Q_2 = 1$ isotropic, and $q^{m-5}Q_{m-4}Q_2/Q_1 = 0$ non-isotropic rays. There are $Q_1/d - 1 = 2$ such sets in H_m.

We obtain the numbers of conjugates in the other 10 sets by determining for each set the subgroup permutable with one of its matrices. To save space, we let I, P, A, B denote respectively the identity, a matrix of order p, and two arbitrary matrices, each on 2 or $m - 2$ variables, and let θ and ϕ be multipliers $\neq 1$ such that $\theta\bar{\theta} = \phi\bar{\phi} = 1$. Then by $[\theta, \bar{\theta}, P]$, for instance, we mean a matrix of order pQ_1 which multiplies the first two variables respectively by θ and $\bar{\theta}$, and transforms the remaining $m - 2$ variables by an operation of order p. Using this notation, we obtain the sets of conjugates as follows.

(11) The multiplication $[\theta, \bar{\theta}, I]$ is permutable only with matrices of the form $[\phi^\alpha, \phi^\beta, A]$ when $\theta \neq \pm 1$. There are $Q_1^2g_{m-2}$ of these in G_m, corresponding to $Q_1^2h_{m-2}$ in H_m, if g_m and h_m denote the respective orders of G_m and H_m. Hence the matrix $[\theta, \bar{\theta}, I]$ belongs to a set of $h_m/h_{m-2}Q_1^2 = q^{2m-3}Q_mQ_{m-1}/Q_1^2 = 480$ conjugate matrices. It leaves $Q_{m-2}Q_{m-3}/Q_2 = 3$ isotropic rays invariant, multiplies two non-isotropic rays respectively by θ and $\bar{\theta}$, and leaves $q^{m-3}Q_{m-2}/Q_1 = 2$ others invariant. Using the various possible values of θ, we have altogether $[q/2] = 1$ such set. For q odd, we should also have a set for $\theta = -1$, consisting of $h_m/h_{m-2}g_2$ matrices, whose characters in the representations we are studying would be $\pm Q_1 + Q_{m-2}Q_{m-3}/Q_2$ and $\pm(q^2 - q) + q^{m-3}Q_{m-2}/Q_1$ respectively.

(12) The matrix $[\theta, \bar{\theta}, P]$ is permutable only with $Q_1^3g_{m-2}/Q_{m-2}Q_{m-3} = qQ_1^2 = 18$ matrices, and belongs to a set of $q^{2m-3}Q_mQ_{m-1}Q_{m-2}Q_{m-3}/Q_1^3 = q^5Q_4Q_3Q_2/Q_1^2 = 1440$ conjugates. It leaves one isotropic ray invariant, and two non-isotropic rays relatively invariant with factors θ and $\bar{\theta}$ respectively. There is $[q/2] = 1$ such set in H_m.

(13) The multiplication $[\theta I, \bar{\theta} I]$ is permutable only with matrices of the form $[A, B]$. There are $g_{m-2}g_2Q_1/d = q^2Q_2^2Q_1/d = 108$ of these in H_m. Hence there are $h_md/g_{m-2}g_2Q_1 = q^{2m-4}Q_mQ_{m-1}/Q_2Q_1 = 240$ matrices in this set of conjugates. Q_1 isotropic rays and $q^2 - q$ non-isotropic rays have the multiplier θ, and (when $m = 4$) an equal number have the multiplier $\bar{\theta}$. There is $[q/2] = 1$ such set in H_m.

(14) The matrix $[\theta P, \bar{\theta} P]$ is permutable only with $(qQ_1/d)g_{m-2}Q_1/Q_{m-2}Q_{m-3} =$

12 matrices of H_m. It belongs to a set of $q^{2m-4}Q_mQ_{m-1}Q_{m-2}Q_{m-3}/Q_1^2 = 2160$ conjugate matrices whose periods divide pQ_1. Two isotropic rays are relatively invariant, with the respective multipliers θ and $\bar{\theta}$. No non-isotropic rays are invariant (for $m = 4$). There is $[q/2] = 1$ such set in H_m.

(15, 16) The matrix $[\theta P, \bar{\theta} I]$ is permutable with just $g_{m-2}qQ_1/d$ matrices of H_m. Hence there are $q^{2m-4}Q_mQ_{m-1}/Q_1 = 720$ matrices in the corresponding set of conjugates. One isotropic ray is relatively invariant with multiplier θ, and Q_1 with multiplier $\bar{\theta}$. Also $q^2 - q$ non-isotropic rays are multiplied by $\bar{\theta}$. There are $Q_1/d_2 - 1 = 2$ such sets in H_m, if d_2 is the h.c.f. of Q_1 and 2.

(17) The cyclic permutation T: $(x_1x_2x_3x_4)$ is permutable with the subgroup H_4 of order 8 generated by itself and the matrix $(1 - \delta_{ij})$. It leaves invariant one isotropic ray and no others, and belongs to a set of 3240 conjugates.

(18) The matrix

$$S:\quad \begin{pmatrix} 0 & \omega^2 & \omega^2 & \omega^2 \\ 0 & \omega & \omega^2 & 1 \\ 0 & 1 & \omega^2 & \omega \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

is of order 5, and can be shown to be permutable only with its own powers. It leaves no rays invariant, and belongs to a set of 5184 conjugates. *The matrices S and T^2 generate the entire group $HO(4, 4)$ of order* 25920.[4]

(19, 20) The matrix ST^2 is of order 9, is permutable only with its own powers, and is therefore one of 2880 conjugate matrices. It leaves no isotropic rays invariant, but multiplies the vector $(1, 0, \omega, \omega^2)$ by $\omega(= \theta)$. Its inverse belongs to another conjugate set with multiplier ω^2 $(= \bar{\theta})$.

We now have accounted for all the 25920 matrices of $HO(4, 4)$ in 20 sets of conjugates, and have found the characters of six reducible representations of degrees 45 and 40. They were seen to be the following, where θ is in turn each of the three cube roots of unity, 1, ω, and ω^2:

$$45,\ 13,\ 5,\ 1,\ 9\theta,\ 9\bar{\theta},\ \theta,\ \bar{\theta},\ \theta,\ \bar{\theta},\ 3,\ 1,\ 3\theta + 3\bar{\theta},\ \theta + \bar{\theta},\ 3\theta + \bar{\theta},\ \theta + 3\bar{\theta},\ 1,\ 0,\ 0,\ 0;$$
$$40, 8, 0, 4, 1 + 12\theta, 1 + 12\bar{\theta}, 1 + 4\theta, 1 + 4\bar{\theta}, 1, 1, 2 + \theta + \bar{\theta}, \theta + \bar{\theta}, 2\theta + 2\bar{\theta}, 0, 2\theta, 2\bar{\theta}, 0, 0, \theta, \bar{\theta}.$$

These six representations can be split into nine distinct irreducible components by applying the theory developed in §2. The two representations corresponding to permutation groups have in common the identity representation and one other irreducible representation of degree $q^3Q_{m-1}Q_{m-2}/Q_2Q_1 = 24$, and their remaining irreducible components are of degree 20 and 15 respectively. The four monomial representations are conjugate imaginary in pairs. Those of degree 45 have a common real component of degree $Q_mQ_{m-3}/Q_2 = 15$, and the remaining components of degree 30 are also found as components in the monomial representations of degree 40, whose second components are of degree 10.

[4] H. R. Brahana has already given two such generators for the simply isomorphic group $A(4, 3)$. See Annals of Mathematics, vol. 31 (1930), p. 533.

h_λ																					h/h_λ
1	1	24	20	15	15	30	30	10	10	64	81	6	5	5	40	40	45	45	30	60	25920
45	1	8	4	-1	7	6	6	2	2	0	9	-2	-3	-3	-8	-8	-3	-3	-10	-4	576
270	1	0	4	-1	3	2	2	-2	-2	0	-3	2	1	1	0	0	-3	-3	2	4	96
540	1	0	0	3	-1	2	2	2	2	0	-3	2	1	1	0	0	1	1	-2	0	48
40	1	6	2	6	-3	$9\omega+3$	$9\omega^2+3$	$3\omega-2$	$3\omega^2-2$	-8	0	-3	$3\omega+2$	$3\omega^2+2$	$6\omega^2-2$	$6\omega-2$	-9ω	$-9\omega^2$	3	6	648
40	1	6	2	6	-3	$9\omega^2+3$	$9\omega+3$	$3\omega^2-2$	$3\omega-2$	-8	0	-3	$3\omega^2+2$	$3\omega+2$	$6\omega-2$	$6\omega^2-2$	$-9\omega^2$	-9ω	3	6	648
360	1	2	-2	2	1	$\omega-1$	ω^2-1	$3\omega+2$	$3\omega^2+2$	0	0	1	ω^2-1	$\omega-1$	-2ω	$-2\omega^2$	3ω	$3\omega^2$	-1	2	72
360	1	2	-2	2	1	ω^2-1	$\omega-1$	$3\omega^2+2$	$3\omega+2$	0	0	1	$\omega-1$	ω^2-1	$-2\omega^2$	-2ω	$3\omega^2$	3ω	-1	2	72
2160	1	0	0	0	-1	$-\omega^2$	$-\omega$	$-\omega$	$-\omega^2$	0	0	-1	ω	ω^2	0	0	ω	ω^2	1	0	12
2160	1	0	0	0	-1	$-\omega$	$-\omega^2$	$-\omega^2$	$-\omega$	0	0	-1	ω^2	ω	0	0	ω^2	ω	1	0	12
480	1	3	-1	0	3	0	0	1	1	-2	0	0	2	2	1	1	0	0	3	-3	54
1440	1	-1	1	2	1	0	0	-1	-1	0	0	-2	0	0	1	1	0	0	-1	-1	18
240	1	0	5	3	0	-3	-3	1	1	4	0	3	-1	-1	-2	-2	0	0	3	-3	108
2160	1	0	1	-1	0	-1	-1	1	1	0	0	-1	1	1	0	0	0	0	-1	1	12
720	1	2	1	-1	-2	$2\omega+1$	$2\omega^2+1$	-1	-1	0	0	1	$2\omega^2+1$	$2\omega+1$	$-2\omega^2$	-2ω	0	0	-1	-1	36
720	1	2	1	-1	-2	$2\omega^2+1$	$2\omega+1$	-1	-1	0	0	1	$2\omega+1$	$2\omega^2+1$	-2ω	$-2\omega^2$	0	0	-1	-1	36
3240	1	0	0	-1	1	0	0	0	0	0	-1	0	-1	-1	0	0	1	1	0	0	8
5184	1	-1	0	0	0	0	0	0	0	-1	1	1	0	0	0	0	0	0	0	0	5
2880	1	0	-1	0	0	0	0	ω	ω^2	1	0	0	$-\omega^2$	$-\omega$	ω	ω^2	0	0	0	0	9
2880	1	0	-1	0	0	0	0	ω^2	ω	1	0	0	$-\omega$	$-\omega^2$	ω^2	ω	0	0	0	0	9

The arithmetical properties of tables of characters now enable us to determine the characters of all the irreducible representations. For instance, the conjugate set (18) can have only five characters different from zero—rational integers the sum of whose squares is 5—which can only be ± 1. Hence only these five of the 20 irreducible representations can have degrees not divisible by 5. If these be 1, 24, x, y, z, we shall have $1 - 24 \pm x \pm y \pm z = 0$, where $\pm x \equiv \pm y \equiv \pm z \equiv 1 \pmod 5$, and x, y, z each divide 25920. The choice $x = 4$, $y = 54$, $z = 81$ leads to incompatible conditions for the characters in the conjugate set (2); the choices $x = 9$, $y = 4$, 16, or 64, $z = 36$, 16, or 96 do likewise for the characters in the conjugate set (5). There remains only the choice $x = 64$, $y = 81$, $z = 6$. Since there are no other representations of these degrees, the characters must be integers, and are found to be uniquely determined from the table. The degrees of the remaining 8 irreducible representations are all divisible by 5, and it can be shown that half are divisible by 2, half by 3, and an odd number by 4. Since the sum of the squares of these degrees is 11800, they can be determined uniquely: 5, 5, 40, 40, 45, 45, 30, 60. After some further computations of a similar nature, the complete table of characters can be filled in, as shown on page 483. Each row corresponds to a complete set of h_λ conjugates, and each column to an irreducible representation. The numbers h_λ are given on the left, and h/h_λ on the right.

Brown University.

ABEL-POISSON SUMMABILITY OF DERIVED CONJUGATE FOURIER SERIES

By A. F. Moursund

1. **Introduction.**[1] In this paper we give theorems concerning the Abel-Poisson summability of the r-th, $r = 0, 1, 2, \cdots$, derived series of the conjugate series of the Fourier series.[2] These theorems may be considered as extensions of theorems given by B. N. Prasad,[3] A. Plessner,[4] and others for the summability of the conjugate series and its first derived series.

2. **Notation.** Throughout this note we assume that the function $f(x)$ is Lebesgue integrable on $(-\pi, \pi)$ and of period 2π. The letters r and p always represent positive integers or zero, and the letter K represents a positive absolute constant which need not be always the same even in a single discussion. For convenience we designate a fixed value of x by x instead of the usual x_0.

We set

$$p(v, s) \equiv (1 + s^2 - 2s \cos v)^{-1}, \qquad P(v, s) \equiv sp(v, s) \sin v; \tag{1}$$

$$M_r^{(p)}(v, s) \equiv v^{r+p+1} \frac{\partial^{r+p+1}}{\partial v^{r+p+1}} \{P(v, s) - 1/2 \cot v/2\}; \tag{2}$$

$$V(s, x) \equiv \sum_{n=1}^{\infty} (-b_n \cos nx + a_n \sin nx)s^n \qquad (0 \leqq s < 1), \tag{3}$$

Received September 5, 1935; presented to the American Mathematical Society, September 13, 1935.

[1] The author is indebted to the referee for many helpful suggestions.

[2] No theorems concerning the Abel-Poisson summability of the r-th, $r > 1$, derived conjugate series appear in the literature. For theorems concerning the summability of such series by other methods see the following papers: A. F. Moursund, *On summation of derived series of the conjugate Fourier series*, Annals of Mathematics, (2), vol. 36 (1935), pp. 182–193, and American Journal of Mathematics, vol. 57 (1935), pp. 854–860; A. H. Smith, *On the summability of derived conjugate series of the Fourier-Lebesgue type*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 406–412; A. F. Moursund, *On the r-th derived conjugate function*, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 131–136. Since Abel-Poisson summability follows from Cesàro summability it is possible to obtain theorems which are similar to, but less general than, some of the theorems of this note from Theorem 6.4 of the second paper cited here. Theorem 1 does not follow from any of the results of papers listed here.

[3] B. N. Prasad, *Contribution à l'étude de la série conjuguée d'une série de Fourier*, Journal de Mathématiques Pures et Appliquées, (9), vol. 11 (1932), pp. 153–205. This paper gives a long bibliography.

[4] A. Plessner, *Zur Theorie der konjugierten trigonometrischen Reihen*, Mitteilungen des Mathematischen Seminars der Universität Giessen, Heft 10 (1923).

where a_n and b_n are the Fourier coefficients of $f(x)$;

$$(4)\quad \begin{aligned} A_r^{(0)}(v) &\equiv f(x+v) + (-1)^{r+1} f(x-v) - 2 \sum_{i=0}^{[(r-1)/2]} \alpha_{r-1-2i}\, v^{r-1-2i}, \\ A_r^{(p+1)}(v) &\equiv 1/p! \int_0^v (v-t)^p A_r^{(0)}(t)\, dt, \end{aligned}$$

where $\alpha_{r-1}, \alpha_{r-2}, \cdots, \alpha_{r-1-2i}$, $0 \leqq 2i \leqq r-1$, are arbitrary numbers;[5]

$$(5)\quad \begin{aligned} C_r &\equiv \sum_{i=0}^{[r/2]-1} \alpha_{r-1-2i}\,(r-1-2i)! \sum_{j=i}^{[r/2]-1} \frac{\pi^{2j-2i}}{(2j+1-2i)!} \frac{d^{2j+1}}{dv^{2j+1}} \cot v/2 \Big]_{v=\pi}, \\ C_r^{(p)} &\equiv -C_r + 1/2\pi \sum_{j=[r/2]}^{[(r+p)/2]-1} A_r^{(2j-r+2)}(v) \frac{d^{2j+1}}{dv^{2j+1}} \cot v/2 \Big]_{v=\pi}. \end{aligned}$$

In (4), (5), and elsewhere in this note, a summation is to be interpreted as zero whenever the upper limit is less than the lower limit. Thus C_0, C_1, and $C_r^{(0)}$ are equal to zero, and

$$(6)\quad A_0^{(0)}(v) \equiv f(x+v) - f(x-v), \qquad A_1^{(0)}(v) \equiv f(x+v) + f(x-v) - 2\alpha_0.$$

We find it convenient to set

$$(7)\quad \epsilon \equiv \arcsin(1-s) \qquad (0 \leqq s < 1),$$

and write $o(1)$ as $\epsilon \to 0$ for $o(1)$ as $s \to 1$.

3. **Results.** Our principal results are given by the following theorems.

THEOREM 1. *If*

$$(8)\quad A_r^{(p+1)}(v) = o(v^{r+p+1}), \qquad \text{as } v \to +0,$$

then

$$(9)\quad \frac{\partial^r}{\partial x^r} V(s,x) + \frac{(-1)^{r+p}}{2\pi} \int_\epsilon^\pi A_r^{(p)}(v) \frac{d^{r+p}}{dv^{r+p}} \cot v/2\, dv \to C_r^{(p)}, \qquad \text{as } \epsilon \to 0.\text{[6]}$$

[5] An equivalent definition for $A_r^{(p+1)}(v)$ is given by

$$A_r^{(p+1)}(v) \equiv \int_0^v A_r^{(p)}(t)\, dt \qquad (p = 0, 1, 2, \cdots),$$

i.e., $A_r^{(p+1)}(v)$ is the $(p+1)$-fold integral of $A_r^{(0)}(v)$. For either definition $A_r^{(p+1)}(v)$ is the q-fold integral of $A_r^{(p+1-q)}(v)$.

[6] The condition $A_r^{(p+1)}(v) = o(v^{r+p+1})$ is equivalent to the condition

$$\int_0^v A_r^{(p)}(t)/t^{r+p}\, dt = o(v).$$

See Smith, loc. cit., pp. 411–412, for a proof of the equivalence of two similarly related conditions. By setting $r = 0$, $p = 1$, in Theorem 1 we obtain Prasad's Theorem 3 (loc. cit., p. 173) for the summability of the conjugate series; and by setting $r = 1$, $p = 0$, we obtain Plessner's Theorem 3 (loc. cit., p. 4) for the summability of the first derived series of the conjugate series. Plessner's Theorem 1 for the conjugate series is included by Prasad's theorem.

THEOREM 2. *If*

$$\lim_{\delta\to+0}\int_\delta^\pi A_r^{(p)}(v)\,\frac{d^{r+p}}{dv^{r+p}}\cot v/2\,dv \tag{10}$$

exists, then (8) *holds and accordingly*

$$\frac{\partial^r}{\partial x^r}V(s,x)\to\frac{(-1)^{r+p+1}}{2\pi}\lim_{\delta\to+0}\int_\delta^\pi A_r^{(p)}(v)\,\frac{d^{r+p}}{dv^{r+p}}\cot v/2\,dv+C_r^{(p)}, \tag{11}$$

as $s\to1-0$.

THEOREM 3. *If* $r\geqq1$ *and* $d^{r-1}f(x)/dx^{r-1}$ *is of bounded variation, then* (10) *and consequently* (9) *and* (11) *hold for every* p *and almost all* x *with*

$$\alpha_{r-1-2i}\equiv\frac{d^{r-1-2i}f(x)/dx^{r-1-2i}}{(r-1-2i)!}.$$

THEOREM 4. *If* $r\geqq0$ *and* $f^{(r+1)}(x)$, *the generalized derivative of order* $r+1$ *of* $f(x)$, *exists at the point* x [*i.e., if* $f(x)$ *satisfies an equation of the form*

$$\frac{f(x+v)+(-1)^{r+1}f(x-v)}{2}$$

$$=\sum_{i=0}^{[(r+1)/2]}f^{(r+1-2i)}(x)\,\frac{v^{r+1-2i}}{(r+1-2i)!}+w(x,s)\,\frac{v^{r+1}}{(r+1)!}.$$

where $w(x,v)\to0$, *as* $v\to+0$], *then* (10) *and consequently* (9) *and* (11) *hold for every* p *with* $\alpha_{r-1-2i}\equiv f^{(r-1-2i)}(x)/(r-1-2i)!$.

4. **An expression for** $\partial^rV(s,x)/\partial x^r$. In this section we obtain the expression for $\partial^rV(s,x)/\partial x^r$ used in the proof of our theorems.

LEMMA 1. *For* $0\leqq s<1$

$$P(v,s)=\sum_{n=1}^\infty s^n\sin nv;\qquad \frac{\partial^{2k}}{\partial v^{2k}}P(v,s)=0\Big]_{v=0,\,\pi};$$

and

$$\frac{\partial^{2k+1}}{\partial v^{2k+1}}P(v,s)\Big]_{v=\pi}=\tfrac12\frac{\partial^{2k+1}}{\partial v^{2k+1}}\cot v/2\Big]_{v=\pi}+o(1),\qquad as\ s\to1.$$

Proof.[7] For $v\neq0$

$$\frac{\partial^{2k+1}}{\partial v^{2k+1}}\{P(v,s)-\tfrac12\cot v/2\}$$

$$=-\frac{(1-s)^2}{2}\frac{\partial^{2k+1}}{\partial v^{2k+1}}\{p(v,s)\cot v/2\}=o(1),\qquad \text{as } s\to1.$$

[7] For a proof of the first part of this lemma see T. J. I'a. Bromwich, *Theory of Infinite Series*, 1926, pp. 186–187. The second part follows from the first part, so we prove only the third part.

LEMMA 2. *For* $r \geqq 0$

$$(12)\qquad \frac{2(-1)^{r+1}}{\pi}\int_0^\pi \sum_{i=0}^{[(r-1)/2]} \alpha_{r-1-2i}\, v^{r-1-2i}\,\frac{\partial^r}{\partial v^r}P(v,s)\,dv = -C_r + o(1),$$

as $s \to 1$.

Proof. Integrating by parts and using Lemma 1, we have

$$\int_0^\pi v^{r-1-2i}\frac{\partial^r}{\partial v^r}P(v,s)\,dv = \sum_{j=1}^{r-2i-1}(-1)^{j-1}\frac{\partial^{r-j}}{\partial v^{r-j}}P(v,s)\frac{(r-1-2i)!}{(r-2i-j)!}v^{r-2i-j}\Big]_0^\pi$$

$$= \sum_{k=i}^{[r/2]-1}(-1)^r\frac{\partial^{2k+1}}{\partial v^{2k+1}}P(v,s)\Big]_{v=\pi}\cdot\frac{(r-1-2i)!}{(2k-2i+1)!}\pi^{2k-2i+1}$$

$$= \frac{1}{2}\sum_{k=i}^{[r/2]-1}\frac{(-1)^r(r-1-2i)!}{(2k-2i+1)!}\pi^{2k-2i+1}\frac{\partial^{2k+1}}{\partial v^{2k+1}}\cot v/2\Big]_{v=\pi} + o(1)\quad \text{as } s\to 1.$$

The lemma follows when we multiply by the proper constant factor and sum with respect to i between $i = 0$ and $[(r-1)/2]$. The term which is not in C_r introduced when r is odd vanishes.

LEMMA 3. *For* $r \geqq 0$

$$(13)\qquad \frac{\partial^r}{\partial x^r}V(s,x) = \frac{(-1)^{r+p+1}}{\pi}\int_0^\pi A_r^{(p)}(v)\frac{\partial^{r+p}}{\partial v^{r+p}}P(v,s)\,dv + C_r^{(p)} + o(1)\quad \textit{as } s\to 1.$$

Proof. Upon differentiating (3) r times, taking into account the fact that $f(x)$ is of period 2π and simplifying, subtracting the left member and adding the right hand member of (12) (Lemma 2), integrating by parts p times, and using Lemma 1 we have

$$\frac{\partial^r}{\partial x^r}V(s,x) = \frac{\partial^r}{\partial x^r}\left\{\sum_{n=1}^{\infty}(-b_n\cos nx + a_n\sin nx)s^n\right\}$$

$$= \frac{\partial^r}{\partial x^r}\left\{-\pi^{-1}\int_{-\pi}^{\pi}f(v)P(v-x,s)\,dv\right\} = \frac{(-1)^{r+1}}{\pi}\int_{-\pi}^{\pi}f(x+v)\frac{\partial^r}{\partial v^r}P(v,s)\,dv$$

$$= \frac{(-1)^{r+1}}{\pi}\int_0^\pi\left\{f(x+v) + (-1)^{r+1}f(x-v)\right\}\frac{\partial^r}{\partial v^r}P(v,s)\,dv$$

$$= \frac{(-1)^{r+1}}{\pi}\int_0^\pi A_r^{(0)}(v)\frac{\partial^r}{\partial v^r}P(v,s)\,dv - C_r + o(1)\qquad \text{as } s\to 1$$

$$= \frac{(-1)^{r+1}}{\pi}\left\{\sum_{j=1}^{p}(-1)^{j-1}A_r^{(j)}(v)\frac{\partial^{r+j-1}}{\partial v^{r+j-1}}P(v,s)\Big]_0^\pi\right.$$

$$\left. + (-1)^p\int_0^\pi A_r^{(p)}(v)\frac{\partial^{r+p}}{\partial v^{r+p}}P(v,s)\,dv\right\} - C_r + o(1)\qquad \text{as } s\to 1$$

$$= \frac{(-1)^{r+p+1}}{\pi}\int_0^\pi A_r^{(p)}(v)\frac{\partial^{r+p}}{\partial v^{r+p}}P(v,s)\,dv + C_r^{(p)} + o(1)\qquad \text{as } s\to 1.$$

5. **Final lemmas.** We use the following lemmas in proving our theorems. The reader can readily supply the proofs and parts of proofs which are omitted.

LEMMA 4. *For* $0 < v \leqq \pi$

$$\frac{d^r}{dv^r}\cot v/2 = (-1)^r\left[(1/2)^r r!\cot^{r+1} v/2 + \sum_{i=0}^{[(r-1)/2]} a_i \cot^{r-1-2i} v/2\right],$$

where the a's are positive constants. Hence $\dfrac{d^{2i}}{dv^{2i}}\cot v/2 = 0$ *when* $v = \pi$.

LEMMA 5. *For* $0 \leqq v \leqq \pi$ *and* $0 \leqq s < 1$, *the function* $v^{j+2}\,\partial^j p(v, s)/\partial v^j$ *is uniformly bounded.*

Proof. Obviously

$$v^2\frac{\partial^0}{\partial v^0}p(v, s) = v^2\{(1-s)^2 + 4s\sin^2 v/2\}^{-1} < K.$$

Assuming that the lemma is true for $j = 0, 1, 2, \cdots, n-1$, we have for $j = n$

$$\begin{aligned}
v^{n+2}\left|\frac{\partial^n}{\partial v^n}p(v, s)\right| &= \left|2s\frac{d^{n-1}}{dv^{n-1}}\{\sin v[p(v, s)]^2\}\right| v^{n+2}\\
&\leqq v^{n+2}\sum_{i=0}^{n-1} K\left|\frac{d^i}{dv^i}\{\sin vp(v, s)\}\frac{d^{n-1-i}}{dv^{n-1-i}}p(v, s)\right|\\
&\leqq v^{n+2}\sum_{i=0}^{n-1} K/v^{n+1+i}\left|\frac{d^i}{dv^i}\{\sin v\cdot p(v, s)\}\right|\\
&\leqq v^{n+2}\sum_{i=0}^{n-1} K/v^{n+1-i}\left\{\left|\sin v\frac{d^i}{dv^i}p(v, s)\right| + \sum_{k=1}^{i} K\left|\frac{d^k}{dv^k}\sin v\frac{d^{i-k}}{dv^{i-k}}p(v, s)\right|\right\} < K.
\end{aligned}$$

The lemma follows by mathematical induction.

LEMMA 6. *For* $0 \leqq v \leqq \pi$ *and* $0 < s < 1$, $\epsilon^{r+1}\,\partial^r P(v, s)/\partial v^r$ *is uniformly bounded.*

Proof. Using Lemma 1 we have

$$\begin{aligned}
\epsilon^{r+1}\left|\frac{\partial^r}{\partial v^r}P(v, s)\right| &\leqq \epsilon^{r+1}\sum_{n=1}^{\infty} n^r s^n \leqq \epsilon^{r+1}\sum_{n=1}^{\infty}\frac{(r+n-1)!}{(n-1)!}s^n\\
&= \epsilon^{r+1}\cdot s\cdot r!\,(1-s)^{-(r+1)} \leqq K.
\end{aligned}$$

LEMMA 7. *For* $0 \leqq v \leqq \pi$ *and* $0 \leqq s < 1$, *the function* $v^{r+3}\partial^r\{p(v, s)\cot\frac{1}{2}v\}/\partial v^r$ *is uniformly bounded.*

Proof. Upon differentiating the product by Leibnitz' formula the lemma follows from Lemmas 4 and 5.

LEMMA 8. *For* $0 \leqq v \leqq \pi$ *and* $0 \leqq s < 1$, *the function* $v^3\epsilon^{-2}M_r^{(p)}(v, s)$ *is uniformly bounded.*

Proof. Since

$$M_r^{(p)}(v, s) = -\frac{(1-s)^2}{2}\frac{\partial^{r+p+1}}{\partial v^{r+p+1}}\{p(v, s)\cot v/2\}\cdot v^{r+p+1},$$

the reader can establish the lemma with the aid of Lemma 7.

6. **Proof of Theorem** 1. Upon making use of the expression for

$$\partial^{r+1}V(s, x)/\partial v^{r+1}$$

given by (13) [Lemma 3] we see that proving Theorem 1 is equivalent to showing that when (8) holds

$$\int_0^\pi A_r^{(p)}(v) \frac{\partial^{r+p}}{\partial v^{r+p}} P(v, s)\, dv - \tfrac{1}{2}\int_\epsilon^\pi A_r^{(p)}(v) \frac{d^{r+p}}{dv^{r+p}} \cot \tfrac{1}{2}v\, dv \to 0, \tag{14}$$

as $\epsilon \to 0$.

Upon integrating both of the integrals in (14) by parts and using Lemmas 1 and 6 when $r + p$ is odd and Lemmas 1 and 4 when $r + p$ is even, we have

$$-\int_0^\pi A_r^{(p+1)}(v) \frac{\partial^{r+p+1}}{\partial v^{r+p+1}} P(v, s)\, dv - \tfrac{1}{2}A_r^{(p+1)}(v) \frac{d^{r+p}}{dv^{r+p}} \cot \tfrac{1}{2}v \Big]_{v=\epsilon}$$

$$+ \tfrac{1}{2}\int_0^\pi A_r^{(p+1)}(v) \frac{d^{r+p+1}}{dv^{r+p+1}} \cot \tfrac{1}{2}v\, dv + o(1)$$

$$= -\int_\epsilon^\pi A_r^{(p+1)}(v) \frac{\partial^{r+p+1}}{\partial v^{r+p+1}} \{P(v, s) - \tfrac{1}{2}\cot \tfrac{1}{2}v\}\, dv$$

$$-\int_0^\epsilon A_r^{(p+1)}(v) \frac{\partial^{r+p+1}}{\partial v^{r+p+1}} P(v, s)\, dv \equiv I_1 + I_2 .$$

By Lemma 8

$$|I_1| = \left| \int_\epsilon^\pi \{A_r^{(p+1)}(v)/v^{r+p+1}\} M_r^{(p)}(v, s)\, dv \right|$$

$$\leqq [o(1) \text{ as } \delta \to 0] \int_\epsilon^\delta K\epsilon^2 v^{-3}\, dv + K\int_\delta^\pi \epsilon^2 v^{-3}\, dv$$

$$= [o(1) \text{ as } \delta \to 0] + [o(1) \text{ as } \epsilon \to 0, \text{ for a fixed } \delta];$$

and by Lemma 6

$$|I_2| = \left| \int_0^\epsilon \{A_r^{(p+1)}(v)/v^{r+p+1}\} \left\{ v^{r+p+1} \frac{\partial^{r+p+1}}{\partial v^{r+p+1}} P(v, s) \right\} dv \right|$$

$$= [o(1) \text{ as } \epsilon \to 0] \int_0^\epsilon \epsilon^{-1} \left| \epsilon^{r+p+2} \frac{\partial^{r+p+1}}{\partial v^{r+p+1}} P(v, s) \right| dv = o(1) \qquad \text{as } \epsilon \to 0.$$

Consequently (14) holds and Theorem 1 is established.

7. **Proof of Theorem** 2. To establish Theorem 2, it will be sufficient to show that (10) implies (8), for (8) implies (9), and (9) and (10) imply (11).

When (10) holds

$$A_r^{(p+1)}(v) = \int_0^v \left\{\frac{d^{r+p}}{dv^{r+p}} \cot \tfrac{1}{2}t\right\}^{-1} \left\{A_r^{(p)}(t) \frac{d^{r+p}}{dt^{r+p}} \cot \tfrac{1}{2}t\right\} dt$$

$$= -\left[\left\{\int_t^v A_r^{(p)}(u) \frac{d^{r+p}}{du^{r+p}} \cot \tfrac{1}{2}u \, du\right\}\left\{\frac{d^{r+p}}{dt^{r+p}} \cot \tfrac{1}{2}t\right\}^{-1}\right]_0^v$$

$$+ \int_0^v \left[\left\{\int_t^v A_r^{(p)}(u) \frac{d^{r+p}}{du^{r+p}} \cot \tfrac{1}{2}u \, du\right\} \frac{d}{dt}\left\{\frac{d^{r+p}}{dt^{r+p}} \cot \tfrac{1}{2}t\right\}^{-1}\right] dt$$

$$= o(v^{r+p+1}),$$

as a consequence of Lemma 4; for when (10) holds,

$$\int_t^v A_r^{(p)}(u) \frac{d^{r+p}}{du^{r+p}} \cot \tfrac{1}{2}u \, du \to 0, \qquad \text{as } v, t \to +0.$$

8. **Remarks about Theorems** 3 **and** 4. The existence of the limit in (10) for $p = q$ implies its existence for $p > q$, but the converse is not necessarily true.[8] When the α's are chosen in terms of the generalized derivatives as in Theorem 4, the existence of (10) for $r = q$ implies its existence[9] for $r = q - 2i$, $0 < 2i \leqq q$. The existence almost everywhere, when $d^{r-1}f(x)/dx^{r-1}$ is of bounded variation, of (10) with the α's chosen as in Theorem 3 can be made to depend on Plessner's proof[10] for the case $r = 1$, $p = 0$.

THE UNIVERSITY OF OREGON.

[8] Prasad, loc. cit., discusses the relationship between the cases $r = 0$, $p = 0$ and $r = 0$, $p = 1$. See p. 178.

[9] Moursund, first citation, Theorem 9.2. A proof of the existence of the limit in (10) with $p = 0$ at a point where the $(r + 1)$-th generalized derivative exists is given in the paper here cited.

[10] Plessner, loc. cit. See Moursund, first paper cited, sections 7 and 9; second paper cited, section 5. Footnote 7 of the second paper corrects an error which appears in the first paper.

A CONSTRUCTION FOR PRIME IDEALS AS ABSOLUTE VALUES OF AN ALGEBRAIC FIELD

By Saunders MacLane

1. **Introduction.** The difficulties of actually constructing the prime ideal factors of a rational prime p in an algebraic field have had a considerable influence upon the development of ideal theory. One of the most practical of the methods for this construction consists of three successive "approximations" to the prime factors of p in terms of certain Newton Polygons, similar to the polygons used in the expansion of algebraic functions. This method, due to Ore,[1] is directly applicable in all but certain exceptional cases. The present paper extends the method to all cases by making not three but any number of successive approximations. To formulate this extension simply, it is necessary to replace the prime ideals by certain corresponding "absolute values", which succinctly express the essential properties of the Newton polygons. In terms of these values, the successive approximations are a natural application of a method of finding possible "absolute values" for polynomials.

To introduce these absolute values, consider the ring $\mathfrak{o}$ of all algebraic integers of an algebraic number field, and let $\mathfrak{p}$ be a prime ideal in $\mathfrak{o}$. Since every integer α of the field can be written in the form $(\alpha) = \mathfrak{p}^m \cdot \mathfrak{b}$, where $\mathfrak{b}$ is an ideal prime to $\mathfrak{p}$, we can write the exact exponent m to which $\mathfrak{p}$ divides α as a function $W\alpha = m$. Because of the unique decomposition theorem,

$$(1) \qquad W(\alpha\cdot\beta) = W\alpha + W\beta, \qquad W(\alpha + \beta) \geqq \min(W\alpha, W\beta).$$

Any function $V\alpha$ which has these two properties is called a non-archimedean value or a "Bewertung"[2] of the ring $\mathfrak{o}$, while the particular function W obtained from $\mathfrak{p}$ may be called a $\mathfrak{p}$-adic value. Every value V of $\mathfrak{o}$ is a constant multiple[3] of some $\mathfrak{p}$-adic value W. Hence absolute values can replace prime ideals.

Received February 10, 1936; presented to the American Mathematical Society, December 31, 1935.

[1] O. Ore, *Zur Theorie der algebraischen Körper*, Acta Math., vol. 44 (1924), pp. 219–314; O. Ore, *Newtonsche Polygone in der Theorie der algebraischen Körper*, Math. Annalen, vol. 99 (1928), pp. 84–117. These papers will be cited as Ore I and Ore II, respectively.

[2] W. Krull, *Idealtheorie*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 4, Heft 3. This text, cited henceforth as Krull I, contains further references on absolute values.

[3] E. Artin, *Ueber die Bewertungen algebraischer Zahlkörper*, Jour. für Math., vol. 167 (1932), pp. 157–159. The theorem may be proved thus: Given V, first show that any rational integer $n = 1 + 1 + \cdots + 1$ has a non-negative value and then from (1) that every algebraic integer has a non-negative value. If the value of an ideal $\mathfrak{b}$ be defined as the minimum of $V\alpha$ for $\alpha \in \mathfrak{b}$, then one and only one prime ideal $\mathfrak{p}$ can have a positive value, and V must be $\mathfrak{p}$-adic. A similar theorem holds when $\mathfrak{o}$ is an abstract ring in which the usual prime-ideal decomposition holds. (B. L. van der Waerden, *Moderne Algebra*, vol. 2, §100.)

In the same way every non-archimedean value V_0 of the rational integers is a "p-adic" value for some rational prime p; that is, for any integer a, $V_0 a$ is $m\delta$, where m is the highest power of p dividing a and δ is a constant >0. If $\mathfrak{p}$ is a prime ideal factor of p in an algebraic field, every $\mathfrak{p}$-adic value W, considered only as a value of the rational integers, coincides with one of the p-adic values V_0. Thus W is an "extension" of V_0.

The equivalence of prime ideals to values enables us to state the problem of constructing the prime ideal factors of a rational prime in the following generalized form (with a notation to be used throughout the paper): *Given a field K and a separable extension $K(\theta)$ generated by a root θ of the irreducible polynomial $G(x)$; given also a "discrete" (see §2) value V_0 of K, to construct all extensions W of V_0 to $K(\theta)$.*

This problem will first be reduced in §2 to that of constructing for the ring of polynomials with coefficients in K those values V which are extensions of V_0 and which assign the defining equation $G(x)$ the value $+\infty$. All values of this polynomial ring can be constructed[4] by successive approximations, which consist essentially in determining first the values of the polynomials of lowest degree (in x and in p). The salient features of this method are summarized in §2. Those approximations which can ultimately give G the desired value $+\infty$ we call "approximants" to G (see §3). Each such approximant is itself a value V_k of the polynomial ring and can be constructed from a previous approximant V_{k-1} by using a unique "equivalence" decomposition of $G(x)$ (see §4) and a "Newton polygon" of $G(x)$ with respect to V_{k-1} (see §5). After a finite number of steps (§8) we obtain a set of approximants corresponding to the desired values or prime ideals of $K(\theta)$. The proof of this fact uses the integers of K(§7) and the exponents of prime ideals (§6). The computation of the degrees of prime ideals in §9 yields a constructive proof of the usual relation between degrees and exponents. Finally, the theorems of §10 summarize the results. A comparison with previous methods is also made. We note that some of the concepts resemble those used by Ostrowski[5] and by Deuring and Krull[6] in the (non-constructive) theory of Galois fields with absolute values.

2. Non-finite values of polynomial rings. A non-archimedean exponential *absolute value* of a ring S is a function V, such that, for a in S, Va is a uniquely defined real number or $+\infty$, with the properties

$$(1) \qquad V(ab) = Va + Vb, \qquad V(a + b) \geqq \min(Va, Vb)$$

[4] S. MacLane, *A construction for absolute values in polynomial rings*, to appear in the Trans. Amer. Math. Soc. Cited henceforth as M. All theorems from M required in the sequel will be explicitly stated, so that we refer to M only for certain proofs.

[5] A. Ostrowski, *Untersuchungen zur arithmetischen Theorie der Körper* (Die Theorie der Teilbarkeit in allgemeinen Körpern), Math. Zeit., vol. 39 (1934), pp. 269–404.

[6] M. Deuring, *Verzweigungstheorie bewerteter Körper*, Math. Ann., vol. 105 (1931), pp. 277–307.

W. Krull, *Galoissche Theorie bewerteter Körper*, S. B. München Akad. Wiss., 1930, pp. 225–238.

for all a and b in S. These properties are called the "product" and "triangle" laws respectively. If we exclude the trivial cases when $Va = 0$ for all a or $Va = \infty$ for all a, the laws (1) imply that $V(1) = V(-1) = 0$, that the equality in the triangle law of (1) must hold whenever $Va \neq Vb$, and that $V(0) = +\infty$. Contrary to previous usage, our definition allows elements not 0 to have the value $+\infty$. However, if $Va \neq \infty$ for all $a \neq 0$, we shall call V a *finite* value. Since $V(a^{-1}) = -V(a)$, every value V of a field must be finite. A value V is *discrete* if every Va is an integral multiple of some fixed $\delta > 0$. The original value V_0 of K is discrete by assumption.

Two elements a and b of S are *equivalent* in V if and only if either $V(a - b) > Va = Vb$ or $Va = Vb = \infty$. We write $a \sim b$ for this equivalence. It is a reflexive, symmetric and transitive relation. An element a is *equivalence-divisible* by b in V if and only if there is a c in S such that $a \sim cb$ in V. For this divisibility we write b/a (in V).

A value V of a ring S is an *extension* of a value V_0 of a subring of S if Va and V_0a are identical for all a in the subring. Our original problem can now be reduced to one concerning the polynomial ring $K[x]$, which consists of all polynomials in x with coefficients in K.

THEOREM 2.1. *There is a one-to-one correspondence between the values W of $K(\theta)$ and those values V of $K[x]$ for which $VG(x) = \infty$. Corresponding values V and W are extensions of identical values of K.*

The proof depends on the homomorphism of $K[x]$ to $K(\theta)$. If the value V with $VG(x) = \infty$ is given, two polynomials congruent mod $G(x)$ must have the same value, so that the value W for any $f(\theta)$ can be defined by $Wf(\theta) = Vf(x)$. The same equation serves to define V when W is given.

The method of the paper M for constructing finite values of $K[x]$ applies without essential change for non-finite values. It consists fundamentally in the formation of a sequence of simple values

$$V_1, V_2, V_3, \cdots, V_{k-1}, V_k, \cdots. \tag{2}$$

To obtain any V_k in (2) from the preceding V_{k-1}, we assign a new value μ_k to a suitable polynomial $\phi_k = \phi_k(x)$. The following conditions[7] must hold:

2.21. ϕ_k is *equivalence-irreducible* in V_{k-1}; that is, $\phi_k/(f(x)\cdot g(x))$ in V_{k-1} always implies $\phi_k/f(x)$ or $\phi_k/g(x)$ in V_{k-1};

2.22. ϕ_k is *minimal* in V_{k-1}; that is, $\phi_k/g(x)$ in V_{k-1} always implies that $\deg \phi_k \leqq \deg g(x)$;

2.23. ϕ_k has the leading coefficient 1 and $\deg \phi_k > 0$;

2.24. $\mu_k > V_{k-1}\phi_k$.

When these are true, we call ϕ_k a *key polynomial* and μ_k a *key value* of ϕ_k over

[7] Functions $f(x)$, $g(x)$ or simply f and g, etc., will always represent polynomials in $K[x]$, while $\deg f(x)$ stands for the degree of $f(x)$. If $f = 0$, $\deg f$ is meaningless, and statements about $\deg f$ are taken to be vacuously true.

V_{k-1}. Given such "key" quantities the new value V_k of any polynomial $f(x)$ is determined from V_{k-1} by first finding the *expansion* of $f(x)$

$$(3) \quad f(x) = f_m(x)\phi_k^m + f_{m-1}(x)\phi_k^{m-1} + \cdots + f_0(x), \qquad \deg f_i(x) < \deg \phi_k$$

in powers of $\phi_k(x)$ with coefficients of degree less than that of ϕ_k, then setting

$$(4) \quad V_k f(x) = \min [V_{k-1}f_m + m\mu_k, V_{k-1}f_{m-1} + (m-1)\mu_k, \cdots, V_{k-1}f_0].$$

The so-defined function V_k is always a value of $K[x]$. We say that V_k is obtained by *augmenting* V_{k-1}, and write

$$(5) \qquad V_k = [V_{k-1}, V_k\phi_k = \mu_k].$$

To apply the condition 2.22 it is convenient to note (M, Theorem 9.3):

2.3. The polynomial $f(x)$ with the expansion (3) is minimal in V_k if and only if $f_m(x)$ is a constant from K and $V_k f(x) = V_k(f_m(x)\phi_k^m)$. In particular, the product of two minimal polynomials is itself minimal.

The construction of any value V of $K[x]$ starts with a "first stage" value V_1 which is defined as in equation (4), except that the first key polynomial ϕ_1 is now taken to be x itself and μ_1 is arbitrary, while the value V_{k-1} used for the coefficients f_i, which are now constants, is simply the originally given value V_0 for K. Given such a V_1, new values can now be defined by repeatedly augmenting V_1. A sequence (2) in which each V_i arises by augmenting V_{i-1} with a pair of keys ϕ_i and μ_i from V_{i-1} is called an *augmented sequence*. Each V_k of such a sequence is an *inductive value*, and may be symbolized as

$$(6) \qquad V_k = [V_0, V_1 x = \mu_1, V_2\phi_2 = \mu_2, \cdots, V_k\phi_k = \mu_k].$$

We assume in addition the conditions (M, Definition 6.1)

2.41. $\deg \phi_i \geqq \deg \phi_{i-1}$ $\qquad (i = 2, 3, \cdots);$

2.42. $\phi_i \sim \phi_{i-1}$ in V_{i-1} is false $\qquad (i = 2, 3, \cdots).$

The last key value μ_k may be $+\infty$, but then there is no key over V_k satisfying these conditions, so that no further augmented value is possible. An infinite augmented sequence (2) also gives a limit value, defined by

$$(7) \qquad V_\infty f(x) = \lim_{k\to\infty} V_k f(x) \qquad \text{(for all } f(x)).$$

We will consider only those inductive or limit values which are extensions of the originally given V_0.

To put the values of $K[x]$ in a normal form, we first choose in K a complete set of "representatives" with respect to V_0, such that each element of K is equivalent in V_0 to one and only one representative. If next the coefficients of the expansion (3) are expanded repeatedly with respect to $\phi_{k-1}, \phi_{k-2}, \cdots$, then $f(x)$ is expressed uniquely in the form

$$(8) \qquad f(x) = \sum_j a_j \phi_1^{m_{1j}} \phi_2^{m_{2j}} \cdots \phi_k^{m_{kj}} \qquad (a_j \in K).$$

The exponent m_{ij} is always less than $(\deg \phi_{i+1})/(\deg \phi_i)$, for $i = 1, \cdots, k - 1$ (see M, §16). If all terms in (8) have the same value in V_k, and if each a_j is one of the previously specified representatives, then $f(x)$ is in a sense *homogeneous* in V_k. Any polynomial is equivalent in V_k to a homogeneous polynomial. Henceforth we require in any inductive or limit value (6) that each ϕ_i be homogeneous in the previously constructed V_{i-1}. Then, since the given V_0 is discrete, *every extension of V_0 to $K[x]$ can be uniquely represented as an inductive or limit value* (M, §8, §16).

3. **Approximants to non-finite values.** Our program requires the construction of values V of $K[x]$ for which $VG(x) = \infty$. Any such V can be obtained from a sequence of suitable inductive values V_k. A V_k which might be so used to construct a V with $VG = \infty$ will be called an "approximant", in an explicit sense now to be given. This involves the way in which V_iG increases in a sequence of inductive values V_i, $i = 1, \cdots, k$. This increase is described by M, Theorems 5.1, 6.4, and 6.5, for any $f(x)$ and any $i \neq k$:

3.11. $$V_kf \geqq V_if;$$
3.12. $V_kf > V_if$ if and only if ϕ_{i+1}/f (in V_i);
3.13. $V_k\phi_i = V_i\phi_i$, and $V_kf = V_if$ whenever $\deg f < \deg \phi_{i+1}$.

Further analysis uses the expansion of $G(x)$ in ϕ_k:

$$G(x) = g_m(x)\phi_k^m + g_{m-1}(x)\phi_k^{m-1} + \cdots + g_0(x). \tag{1}$$

Among the exponents j for which $V_kG = V_k(g_j\phi_k^j)$, let α be the largest and β the smallest. The difference $\alpha - \beta$, which depends on both V_k and G, will be called the *projection* of V_k (symbol: proj V_k). One application is

Lemma 3.2. *If proj $V_k = 0$, then no V with $VG > V_kG$ can be obtained by augmenting V_k.*

Proof. The value of each term in (1) is by 3.13 the same in any V as in V_k. By hypothesis there is but one term of minimum value, so that the triangle law (§2, (1)) proves $VG = V(g_\alpha\phi_k^\alpha) = V_kG$. Since we want only those values V_k leading to $VG = \infty$, we are led to

Definition 3.3. *A k-th approximant to $G(x)$ over V_0 is a k-th stage homogeneous finite inductive value of $K[x]$ which is an extension of V_0 and which has a positive projection.*

Lemma 3.4. *If V_k, given as in* §2, (6), *is a k-th approximant to $G(x)$, then so is V_i for $i = 1, \cdots, k - 1$. Furthermore $\phi_k / G(x)$ in V_{k-1} and*

$$V_kG(x) > V_{k-1}G(x) > \cdots > V_1G(x).$$

First note that in the expansion (1) of $G(x)$

$$V_{k-1}G = \min [V_{k-1}(g_m\phi_k^m), V_{k-1}(g_{m-1}\phi_k^{m-1}), \cdots, V_{k-1}g_0], \tag{2}$$

much as in the definition of V_k. For were $V_{k-1}G$ to exceed the indicated minimum, then by the triangle law $V_{k-1}(g_i\phi_k^i)$ would equal this minimum for at least two i's. Were γ the largest such i, then

$$g_\gamma \phi_k^\gamma \sim g_{\gamma-1}\phi_k^{\gamma-1} + \cdots + g_0 \qquad (\text{in } V_{k-1}).$$

Then ϕ_k^γ would be an equivalence-divisor of the polynomial on the right, which is of smaller degree than ϕ_k^γ, a contradiction because ϕ_k and hence ϕ_k^γ is minimal (see §2, 2.3).

By hypothesis proj $V_k > 0$, so that there is an $\alpha > 0$ with $V_kG = V_k(g_\alpha\phi_k^\alpha)$. As $V_{k-1}\phi_k < V_k\phi_k$, we have by (2) and 3.13

$$V_{k-1}G \leqq V_{k-1}(g_\alpha\phi_k^\alpha) < V_k(g_\alpha\phi_k^\alpha) = V_kG.$$

Hence by 3.12 ϕ_k/G in V_{k-1}, and the remaining conclusions follow by Lemma 3.2. Another useful fact is

LEMMA 3.5. *Let $a(x)$ be a minimal polynomial in V_k, and $r(x)$ the remainder of the division of a polynomial $f(x)$ by $a(x)$. Then $V_kr > V_kf$ if and only if $a(x)/f(x)$ in V_k.*

The proof is exactly like that of M, Lemma 4.3.

4. **Unique equivalence-decomposition.** The construction of an approximant V_{k+1} from a given approximant V_k must by Lemma 3.4 use a key polynomial ϕ_{k+1} which is an equivalence factor of $G(x)$. These factors can be found from the unique equivalence-decomposition of $G(x)$, the existence of which will now be established by a modified euclidean algorithm.[8] We first introduce for any V_k an "effective degree" thus: if $f(x)$ is any polynomial, expanded in powers of ϕ_k as in §2, (3), the largest exponent i for which $V_kf = V_k(f_i\phi_k^i)$ is the *effective degree* of f in ϕ_k, and is denoted by $D_\phi f$. Equivalent polynomials have the same effective degree. The proof of the product law (§2, (1)) for any inductive V_k (see M, §4, end) shows that

$$(1) \qquad D_\phi(fg) = D_\phi f + D_\phi g.$$

If we call a polynomial of effective degree zero an *equivalence-unit*, then $e(x)$ is an equivalence unit if and only if there is an "equivalence-reciprocal" $h(x)$ such that $e(x)\cdot h(x) \sim 1$ (in V_k). For if $e(x)$ has such a reciprocal, then (1) proves that $D_\phi e = 0$. Conversely, if $D_\phi e = 0$, then, by definition of D_ϕ, $e(x)$ is equivalent to the last term $e_0(x)$ in the expansion of e in powers of ϕ_k. As deg $e_0 <$ deg ϕ_k, e_0 is prime to ϕ_k, so that there are polynomials $g(x)$ and $h(x)$ with $g(x)\phi_k + h(x)e_0(x) = 1$. Using the minimal property of ϕ_k, we then conclude that $h(x)e(x) \sim 1$.

LEMMA 4.1. *Any polynomial $f(x)$ can be represented in the form $f(x) \sim e(x)\cdot a(x)$, where $e(x)$ is a unit and $a(x)$ is minimal and has the first coefficient 1. In addition, $f(x)$ and $a(x)$ have the same equivalence-divisors.*

[8] A similar algorithm has been used by A. Fraenkel, *Ueber einfache Erweiterungen zerlegbarer Ringe*, Jour. für Math., vol. 151 (1920), pp. 120–166. Compare Ore I, Theorem 6.

Proof. Expand $f(x)$ as in §2, (3), pick out the first term $f_\alpha(x)\phi_k^\alpha$ of minimum value, and find the equivalence-reciprocal $h(x)$ for the equivalence-unit $f_\alpha(x)$. Then expand the polynomial $h(x)\cdot f(x)$ and drop out all terms not of minimum value. There remains an equivalent polynomial $a(x)$, with an expansion beginning with ϕ_k^α. This $a(x)$ is minimal, and we have $f(x) \sim f_\alpha(x)\cdot a(x)$, as required.

To carry out the euclidean algorithm for two polynomials $f(x)$ and $g(x)$ with $D_\phi f \geqq D_\phi g$, write $g(x) \sim e_1(x)a_1(x)$ in accordance with Lemma 4.1 and divide $f(x)$ by $a_1(x)$, getting

$$f(x) = q(x)\cdot a_1(x) + r_2(x) \qquad D_\phi r_2 < D_\phi a_1. \tag{2}$$

If $V_k r_2 > V_k f$, a_1 and hence g is an equivalence-divisor of f. Otherwise, since a_1 is minimal, $V_k r_2 = V_k f$ and all three terms in (2) have the same value. Repeat this process with $a_1(x)$ and $r_2(x) \sim e_2(x)a_2(x)$, etc., until a remainder exceeding the dividend in value is obtained. The preceding remainder $d(x)$ is the greatest common equivalence-divisor of $f(x)$ and $g(x)$. As usual,

$$d(x) \sim s(x)f(x) + t(x)g(x) \qquad (\text{in } V_k) \tag{3}$$

for suitable $s(x)$ and $t(x)$. To establish (3), it is convenient to note that, unless $g(x) \,/\, f(x)$, all the terms in (3) must be of the same value in V_k.

The properties of equivalence-irreducible polynomials are now obtained as usual from (3). A decomposition of any $f(x)$ into such irreducible factors must exist (because of D_ϕ). If we factor out a suitable unit, these irreducible factors can as in Lemma 4.1 be made minimal and hence key polynomials (§2, Conditions 2.21–2.23).

THEOREM 4.2. *In an inductive value V_k every polynomial $f(x)$ has a decomposition*

$$f(x) \sim e(x)\psi_1(x)\psi_2(x)\cdots\psi_t(x) \qquad (\textit{in } V_k), \tag{4}$$

where $e(x)$ is a unit and each $\psi_i(x)$ is a key polynomial. This decomposition is unique, except for the order of the factors and except that $e(x)$ may be replaced by any equivalent unit and $\psi_i(x)$ by any equivalent key.

If we require the factors $\psi_i(x)$ to be homogeneous in V_k (see §2, (8)), they are then unique. Note also that ϕ_k itself may occur as a factor, by

LEMMA 4.3. *In an inductive V_k, ϕ_k is a key polynomial.*

Proof. Since ϕ_k is a key in V_{k-1}, it has the first coefficient 1. Furthermore $D_\phi \phi_k = 1$; hence in any factorization of ϕ_k one factor is a unit, so that ϕ_k is equivalence-irreducible. Finally, ϕ_k is minimal in V_k.

In many cases the construction of the unique equivalence-decomposition (4) for a given polynomial $f(x)$ in a given V_k can be carried out in a finite number of steps.

THEOREM 4.4. *The decomposition* (4) *is constructive when K is the field of rationals.*

The original value V_0 is then associated with a rational prime p, so that

every rational number is equivalent in V_0 to one of the numbers $c \cdot p^m$, $c = 0, 1, \cdots, p-1$; $V_0 p = 1$. Hence the complete set of representatives for V_0 (see §2, end) includes but a finite number of representatives of each possible value[9] m.

There are but a finite number of minimal homogeneous polynomials $b(x)$ of a given degree d and with first coefficient 1. For any such $b(x)$ may be expanded in powers of $x, \phi_2, \cdots, \phi_k$ as in §2, (8) with a highest coefficient 1 of value 0. Because of the homogeneity, this determines the value of every other non-zero coefficient in the expansion. Since these coefficients are representatives, there is but a finite number of possibilities for each coefficient, and hence but a finite number of polynomials $b(x)$.

If $f(x)$ is to be decomposed, write $f(x) \sim e(x)a(x)$ by Lemma 4.2, find all minimal homogeneous polynomials $b(x)$ of degree less than that of $a(x)$ as above and by trial find which products, if any, are equivalent to $a(x)$.

The decomposition (4) can often be constructed by first decomposing the residue-class of $f(x)$ (cf. §9 and M, part II). We can assume that all factors ϕ_k, if any, have already been removed from f. Then $V_k f(x)$ will be in the previous value-group Γ_{k-1} (M, Lemma 9.2), so that there is a unit polynomial $R(x)$ such that $V_k(Rf) = 0$. In the value V_k the residue-class of any polynomial $g(x)$ is denoted by $H_k g$ and is itself a polynomial in a new variable y (M, Theorem 12.1). In particular, $H_k(Rf)$ is a polynomial with a decomposition

$$H_k(Rf) = \alpha_1(y)\,\alpha_2(y) \cdots \alpha_t(y) \tag{5}$$

into irreducible polynomials $\alpha_i(y)$. But there is essentially just one key polynomial $\psi_i(y)$ in V_k with the residue-class $H_k(\psi_i R_i) = \alpha_i$, for a suitable unit R_i (M, Theorem 13.1). Since the residue-class of a product is the product of the residue-classes

$$H_k(Rf) = H_k(R_1\psi_1 R_2\psi_2 \cdots R_t\psi_t),$$

and since polynomials in the same residue-class are congruent,

$$Rf \equiv R_1 R_2 \cdots R_t \psi_1 \psi_2 \cdots \psi_t \qquad (\text{in } V_k).$$

If we multiply by an equivalence-reciprocal of R, we get the decomposition (4). Consequently, (4) can be constructed in this way whenever (5) can be found; that is, whenever polynomials can be constructively factored in the residue-class field of V_0 in K (see §9). In particular, this method applies when K is the field of rationals.

5. The construction of approximants. If

$$G(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0, \tag{1}$$

[9] Theorem 4.4 is true for any K and V_0 with this property.

the key μ_1 of any first approximant $V_1 = [V_0, V_1x = \mu_1]$ must by Definition 3.3 be so chosen that, for suitable $\alpha > \beta$,

$$\alpha\mu_1 + V_0a_\alpha = \beta\mu_1 + V_0a_\beta \leqq i\mu_1 + V_0a_i \quad (i = 0, \cdots, n), \tag{2}$$

where the inequality holds for $i > \alpha$ or $\beta > i$. To interpret this, plot the points $P_i = (n - i, V_0a_i)$ in a cartesian plane. Then (2) states that the line $P_\alpha P_\beta$ has slope μ_1 and that all the points P_i are either above this line or on the line between P_α and P_β. The line segments $P_\alpha P_\beta$ with this property for some μ_1 form a convex broken line stretching from P_n to P_0. This broken line segment is called the *Newton polygon* of the points P_i; it may be characterized as a convex polygon such that each corner is one of the points P_i, while none of the points lie below the polygon. We have shown that each first approximant V_1 corresponds to a side of this polygon of slope $\mu_1 = V_1x$ and of horizontal projection equal to the "projection" of V_1. Hence

$$\sum \text{proj } V_1 = \deg G, \tag{3}$$

the sum being taken over all first approximants V_1.

Next, given any $(k - 1)$-th approximant V_{k-1} we wish to construct all k-th approximants V_k which can be obtained by augmenting V_{k-1}. Consider first the "terminating case" when $G(x)$ is a homogeneous key polynomial[10] over V_{k-1}. Then by Lemma 3.4 the key polynomial ϕ_k must be an equivalence-divisor of the equivalence-irreducible $G(x)$, whence $\phi_k = G$. We obtain no finite approximants, but only the non-finite value $V_k = [V_{k-1}, V_kG(x) = \infty]$, which by Theorem 2.1 corresponds to a value of $K(\theta)$.

Suppose instead that $G(x)$ is not a homogeneous key polynomial over V_{k-1}. Then by Theorem 4.2 and Lemma 4.3

$$G(x) \sim e(x)\phi_{k-1}(x)^{n_0}\psi_1(x)^{n_1} \cdots \psi_t(x)^{n_t} \quad (\text{in } V_{k-1}), \tag{4}$$

where the $\psi_i(x)$ are homogeneous keys over V_{k-1}, all different and different from $G(x)$ and ϕ_{k-1}, while the exponents n_i are all positive, except perhaps for n_0. An augmented V_k must have a key ϕ_k with $\phi_k / G(x)$ in V_{k-1} (Lemma 3.4) and $\phi_k \neq \phi_{k-1}$ (§2, Condition 2.42). Hence ϕ_k is one of $\psi_1, \cdots, \psi_t$.

If one of these factors ψ_i has been selected as ϕ_k, then $G(x)$ has as in §3, (1) an expansion with coefficients $g_i(x)$. To determine the new value $\mu_k = V_k\phi_k$ to be assigned to ϕ_k, we again use a point $Q_i = (m - i, V_{k-1}g_i(x))$ for each term in the expansion and construct the Newton polygon for these points. The requirement that proj $V_k > 0$ again means that μ_k must be the slope of some side of this polygon. An inductive value requires also that $\mu_k > V_{k-1}\phi_k$, so that we use only the *principal part*[11] cf the polygon, composed of those sides of slope $\mu > V_{k-1}\phi_k$.

[10] For convenience, we assume henceforth that the first coefficient in (1) is $a_n = 1$.

[11] In special cases, this has been called a "Hauptpolygon" by Ore (Ore I, p. 229; Ore II, p. 88) and a "verkürztes Polygon" by Rella, *Ordnungsbestimmungen in Integritätsbereichen und Newtonsche Polygone*, Jour. für Math., vol. 158 (1927), pp. 33–48.

THEOREM 5.1. *If V_{k-1} is a $(k-1)$-th approximant in which $G(x)$ is not a homogeneous key, then the k-th approximants which can be derived by augmenting V_{k-1} are all values $V_k = [V_{k-1}, V_k\phi_k = \mu_k]$ in which $\phi_k \neq \phi_{k-1}$ is any one of the keys in the decomposition (4) of $G(x)$, while, for given ϕ_k, μ_k is the slope of any side of the principal Newton polygon of $G(x)$ with respect to ϕ_k and V_{k-1}. Furthermore*

$$(5) \qquad \sum (\text{proj } V_k)\cdot(\deg \phi(V_k)) = (\text{proj } V_{k-1})\cdot(\deg \phi(V_{k-1})),$$

where the sum is taken over all such augmented V_k, and where $\phi(V)$ represents the last key of V. Hence there is at least one approximant V_k from V_{k-1}.

It remains to prove (5). On the left side of (5) suppose first that ϕ_k is the factor ψ_1 in (4), and consider the power $n = n_1$ to which ϕ_k divides G. Since ϕ_k and hence ϕ_k^n is minimal in V_{k-1}, the remainder

$$r(x) = g_{n-1}\phi_k^{n-1} + g_{n-2}\phi_k^{n-2} + \cdots + g_0$$

obtained on dividing G by ϕ_k^n must by Lemma 3.5 have $V_{k-1}r > V_{k-1}G$. Calculation of $V_{k-1}r$ as in §3, (2) gives

$$(6) \qquad \min [V_{k-1}(g_{n-1}\phi_k^{n-1}), \cdots, V_{k-1}g_0] > V_{k-1}(g_n\phi_k^n) = V_{k-1}G,$$

with the equality because n is the largest exponent with ϕ_k^n / G in V_{k-1}. If we set $\nu = V_{k-1}\phi_k$ and use §3, (2), this becomes

$$\begin{aligned} V_{k-1}g_n + n\nu &\leqq V_{k-1}g_j + j\nu && (j = n+1, \cdots, m) \\ &< V_{k-1}g_i + i\nu && (i = 0, \cdots, n-1). \end{aligned}$$

Geometrically, this means that the line L of slope ν through the point Q_n lies above none of the points Q_j and lies below $Q_{n-1}, \cdots, Q_0$. The convex Newton polygon is hence above or on L, so that the principal polygon, containing those sides of slope exceeding ν, consists of the sides joining Q_n to Q_0. The horizontal projection of the principal polygon for $\phi_k = \psi_1$ is therefore $n = n_1$.

However, proj V_k is by definition (§3) the projection of the corresponding side of the principal polygon. Hence a sum taken over those V_k with ψ_1 as the last key gives $\sum$ proj $V_k = n_1$. Similar equations for all ψ_i yield

$$(7) \quad \sum (\text{proj } V_k)\cdot(\deg \phi_k) = n_1 \deg \psi_1 + \cdots + n_t \deg \psi_t = \deg (\psi_1^{n_1} \cdots \psi_t^{n_t}).$$

But $\psi_1^{n_1}\cdots\psi_t^{n_t}$ is minimal, so that its effective and actual degrees in $\phi = \phi_{k-1}$ must agree. Thus

$$(8) \qquad \deg (\psi_1^{n_1} \cdots \psi_t^{n_t}) = D_\phi(\psi_1^{n_1} \cdots \psi_t^{n_t})\cdot(\deg \phi_{k-1}).$$

Because of (4) the effective degree is

$$(9) \qquad D_\phi(\psi_1^{n_1} \cdots \psi_t^{n_t}) = D_\phi G - D_\phi\phi_{k-1}^{n_0} = D_\phi G - n_0.$$

If the expansion of $G(x)$ is $\sum h_i(x)\phi_{k-1}^i$, then $D_\phi G$ is by definition the exponent of the first term of minimum value, while n_0, the highest power with $\phi_{k-1}^{n_0}/G$

in V_{k-1}, is by the argument used in (6) simply the exponent of the last term of minimum value in the expansion of $G(x)$. By the definition of the projection,

$$D_\phi G - n_0 = \text{proj } V_{k-1}. \tag{10}$$

The last four equations combine to give the result (5). By induction on k we obtain from (3) and (5) the following result.

THEOREM 5.2. *If the "terminating" case does not occur by the k-th stage, there is a finite number of k-th approximants, such that* [12]

$$\sum (\text{proj } V_k)\cdot(\deg \phi(V_k)) = \deg G, \tag{11}$$

the sum being taken over all k-th approximants V_k.

THEOREM 5.3. (Terminating case.) *If there is a non-finite homogeneous inductive value V_k with $V_kG = \infty$, then for $i < k$ the value V_i from which V_k is obtained is the only i-th approximant*

Proof. By Lemma 3.2, V_{k-1}, and hence by Lemma 3.4 each V_i, is an approximant. Since $V_kG = \infty$ and G is irreducible, G must be the last key of V_k, whence G is minimal in V_{k-1} (see §2, 2.3):

$$G(x) = \phi_{k-1}^m + g_{m-1}\phi_{k-1}^{m-1} + \cdots + g_0(x).$$

Since G is minimal and (§2, 2.42) cannot be equivalence-divisible by ϕ_{k-1}, the first and last terms here take on the minimum value $V_{k-1}G$, so that proj $V_{k-1} = m$. Thus

$$\deg G = m(\deg \phi_{k-1}) = (\text{proj } V_{k-1})\cdot(\deg \phi_{k-1}),$$

and by (11) V_{k-1} is the only $(k-1)$-th approximant. Hence each V_i is the only i-th approximant.

6. **Exponents for values.** To estimate the growth of μ_k we need "value-groups". If in an algebraic number field the prime ideal $\mathfrak{p}$ is a factor of the rational prime p, and if the corresponding $\mathfrak{p}$-adic value W is an extension of the p-adic value V_0, then the highest power e to which $\mathfrak{p}$ divides p is characterized by $V_0p = e(W\mathfrak{p})$. Hence the group of all numbers used as p-adic values is a subgroup of index e in the group of $\mathfrak{p}$-adic values. For any value V of a ring S, the additive group Γ which contains all real numbers $Vb - Vc$ for b and c in S is called the *value group* of V. This group is cyclic if and only if the value V is discrete (§2). If V is an extension of V_0 to $K[x]$ or to $K(\theta)$, the value group Γ_0 of V_0 must be a subgroup of the value group Γ of V. The order of the factor group Γ/Γ_0 is called the *exponent*[13], exp (V).

Now compute this exponent for an inductive value V_k with a value-group Γ_k. The definition of §2, (4) indicates that every number in Γ_k has the form $\gamma + n\cdot\mu_k$, where n is an integer and γ is in Γ_{k-1}. If we consider only the case when μ_k is commensurable with Γ_{k-1} (by M, Theorem 6.7, this is true when-

[12] An invariant interpretation of (11) will be given in §9.

[13] Similarly defined in Deuring, op. cit., p. 281 and Ostrowski, op. cit., p. 322.

ever V_k can be augmented to some V_{k+1}), there is a unique smallest positive integer τ_k with the property that $\tau_k \mu_k \in \Gamma_{k-1}$. By group theory

$$\text{order } (\Gamma_k/\Gamma_{k-1}) = \tau_k , \tag{1}$$

$$\exp (V_k) = \tau_1 \cdot \tau_2 \cdots \tau_k , \tag{2}$$

where τ_i for $i = 1, \cdots, k$ is similarly defined. The assumption that μ_k is commensurable also proves Γ_k discrete. If $\mu_k = \infty$, the formulas still hold if we take $\tau_k = 1$.

In the course of §8 we shall need an estimate for exp (V_k). Since each key polynomial ϕ_{i+1} is homogeneous (§2) in V_i, any two terms in the expansion of ϕ_{i+1} in powers of ϕ_i must be of equal value, so that this expansion appears as a polynomial in $\phi_i^{\tau_i}$ (M, §11). Consequently deg $\phi_{i+1} \geqq \tau_i(\deg \phi_i)$. Combining these inequalities for all i, we find

$$\deg \phi_k \geqq \tau_1 \tau_2 \cdots \tau_k = \exp (V_{k-1}). \tag{3}$$

7. **Integral key polynomials.** It is often convenient to use keys with "integral" coefficients. Here an *integer*[14] with respect to V_0 is an element $a \in K$ with $V_0 a \geqq 0$. All such integers form a ring, and every element of K is a quotient of two such integers. After the usual transformations we can *assume that $G(x)$ has V_0-integers as coefficients and the first coefficient* 1. The Newton polygon of the first stage then must give a $\mu_1 \geqq 0$, so that $V_k x \geqq 0$ for every approximant.

THEOREM 7.1. *In a homogeneous V_{k+1} with $V_{k+1}x \geqq 0$, we have*

$$0 \leqq \mu_1 < \mu_2 < \cdots < \mu_k < \mu_{k+1} , \tag{1}$$

and the keys ϕ_i are all polynomials in x with V_0-integers as coefficients.

The last key ϕ_{k+1} is minimal (2.3), so has a leading term $\phi_k^{u_k}$ and a homogeneous expansion as in §2, (8):

$$\phi_{k+1} = \phi_k^{u_k} + \sum_j a_j \phi_1^{m_{1j}} \phi_2^{m_{2j}} \cdots \phi_k^{m_{kj}} \qquad (a_j \in K, m_{kj} < u_k), \tag{2}$$

where, if n_i stands for deg ϕ_i, the degrees m_{ij} are limited by

$$m_{ij} < n_{i+1}/n_i \qquad (\text{all } j, i = 1, 2, \cdots, k-1). \tag{3}$$

Since ϕ_{k+1} is homogeneous, all terms in (2) have the same value. Hence

$$\mu_{k+1} > V_k \phi_{k+1} = u_k \mu_k = (n_{k+1}\mu_k)/n_k . \tag{4}$$

Since $\mu_1 \geqq 0$, (4) for every k gives (1). We next estimate the terms of (2).

LEMMA 7.2. *In any V_k with $V_k x \geqq 0$, a term*

$$T = \phi_1^{m_1} \phi_2^{m_2} \cdots \phi_{k-1}^{m_{k-1}}, \qquad (m_i < n_{i+1}/n_i \textit{ for all } i)$$

has a value $V_k T \leqq V_k \phi_k$.

[14] Cf. Ostrowski, op. cit., p. 288, or the "Bewertungsring" in Krull, *Idealtheorie*, p. 101.

This inequality can also be written as

$$m_1\mu_1 + \cdots + m_{k-1}\mu_{k-1} \leqq \mu_k.$$

It is true for $k = 1$ or 2, by hypothesis and (4). If we assume it for $k - 1$, then, since n_k/n_{k-1} is integral,

$$\sum_{i=1}^{k-1} m_i\mu_i = m_{k-1}\mu_{k-1} + \sum_{i=1}^{k-2} m_i\mu_i \leqq (m_{k-1} + 1)\mu_{k-1} \leqq \frac{n_k}{n_{k-1}}\mu_{k-1} < \mu_k.$$

Theorem 7.1 now follows by induction. It is true for $k = 1$. If all the keys of V_k have V_0-integral coefficients, all terms in the expansion (2) of ϕ_{k+1} have the same value. But $\phi_1^{m_{1j}} \cdots \phi_k^{m_{kj}} = T \cdot \phi_k^{m_{kj}}$ has by the lemma a value not exceeding $V_k\phi_k^{m_{kj}+1} = V_k\phi_k^{n_k}$. Hence the coefficient a_j has a non-negative value, and a_j is V_0-integral.

Note. If K is the field of rational numbers, $G(x)$ with leading coefficient 1 can be so chosen that all its coefficients are ordinary integers (with non-negative value in every V_0). The same proof then shows that all ϕ_k have ordinary integers as coefficients, provided only that the representatives (§2) for each p-adic value V_0 are chosen as the numbers $c \cdot p^m$, $c = 0, \cdots, p - 1$. Similar results hold when K is an algebraic number field.

8. **The finiteness theorem.** Each k-th approximant may give rise to one or more $(k + 1)$-th approximants, so that the number of k-th approximants can increase with k. Ultimately, the number of approximants stops increasing, but for a finite construction we must be able to tell how soon this is the case:

THEOREM 8.1. *One can find an integer k' so large that each k'-th approximant has the projection 1. As a result, only one $(k + 1)$-th approximant can be obtained by augmenting any given k-th approximant, for any $k \geqq k'$.*

The second conclusion follows from the first, because in §5, (5), deg ϕ_k cannot decrease (§2, Condition 2.41). To establish the first conclusion, we will show that a projection not 1 gives G a multiple factor "mod μ_k", in the sense in which $h(x)$ is a common factor "mod ν" in

LEMMA 8.2. *If, in any homogeneous V_k with $V_k x \geqq 0$, $f(x)$ and $g(x)$ are polynomials with V_0-integral coefficients and a resultant $R(f, g)$, if there are polynomials $h(x)$, $a(x)$, and $b(x)$ with*

$$V_k(f - ha) \geqq \nu,\ V_k(g - hb) \geqq \nu \qquad (\nu \text{ real}),$$

and if $h(x)$ is not a unit in V_k, then $V_k[R(f, g)] \geqq \nu$.

Proof. Since $R(f, g) = 0$ would imply $V_k R = \infty$, we can assume $R(f, g) \neq 0$, so that there exist $c(x)$ and $d(x)$, with V_0-integral coefficients, such that

$$c(x)f(x) + d(x)g(x) = R(f, g).$$

(van der Waerden, *Moderne Algebra*, vol. 2, p. 4). Hence

$$R(f, g) = (ca + db)h + c(f - ha) + d(g - hb).$$

Since $V_k x \geqq 0$ and therefore $V_k c \geqq 0$ and $V_k d \geqq 0$, the last two terms here have values not less than ν. Were $V_k R < \nu$, we should have

$$R(f, g) \sim (ca + db)h \qquad \text{(in } V_k).$$

Since R is a constant, this makes h a unit (see §4), contrary to hypothesis. To apply this lemma when R is a discriminant, use

LEMMA 8.3. *In any homogeneous V_k with $V_k x \geqq 0$ and $V_k \phi_k = \mu_k$ the derivative $f'(x)$ of any polynomial $f(x)$ has a value $V_k f'(x) \geqq V_k f(x) - \mu_k$.*

For $k = 1$ the result follows readily, since the value of a natural integer $1 + \cdots + 1$ is never negative. If the lemma is true for V_{k-1}, and if $f(x)$ has the expansion $\sum f_j(x)\phi_k^j$ as in §2, (3), then

$$f'(x) = \sum f_j'(x)\phi_k^j + \sum jf_j(x)\phi_k^{j-1}\phi_k'(x).$$

The value of the first sum exceeds $V_k f - \mu_k$ because of the induction assumption and because $\mu_k > \mu_{k-1}$. The value of the second sum is $\geqq V_k f - \mu_k$, since $V_k j \geqq 0$ and $V_k \phi_k' \geqq 0$, the latter because ϕ_k has V_0-integral coefficients by Theorem 7.1.

To establish Theorem 8.1, consider a V_k with a projection $\alpha - \beta > 1$. The expansion of §3, (1), used to define this projection gives

$$(1) \qquad V_{k-1}g_\alpha + \alpha\mu_k \leqq V_{k-1}g_i + i\mu_k \qquad (i = 0, \cdots, m).$$

Division of $G(x)$ by ϕ_k^α yields, in terms of this expansion,

$$(2) \qquad G(x) = q(x)\phi_k^\alpha + r(x), \qquad r(x) = \sum_{i=0}^{\alpha-1} g_i(x)\phi_k^i.$$

For this remainder $r(x)$ the triangle law (§2, (1)) and (1) show

$$V_{k-1}r \geqq \min_i [V_{k-1}g_i + i\cdot V_{k-1}\phi_k] \geqq \min_i [V_{k-1}g_\alpha + (\alpha - i)\mu_k + i\cdot V_{k-1}\phi_k],$$

where i ranges from 0 to $\alpha - 1$. Since $\mu_k > V_{k-1}\phi_k$, the minimum is at $i = \alpha - 1$:

$$(3) \qquad V_{k-1}r \geqq V_{k-1}g_\alpha + \mu_k + (\alpha - 1)V_k\phi_k.$$

As the divisor ϕ_k^α has V_0-integral coefficients and first coefficient 1, the quotient and $g_\alpha(x)$ likewise have integral coefficients, whence $V_{k-1}g_\alpha \geqq 0$, since $V_{k-1}x \geqq 0$. Further, (4) of §7 proves $V_{k-1}\phi_k \geqq \mu_{k-1}$, while $\alpha \geqq$ proj V_k was assumed to exceed 1, so that (3) becomes

$$(4) \qquad V_{k-1}r \geqq \mu_k + \mu_{k-1}.$$

Differentiation of (2), with Lemma 8.3, now proves

$$V_{k-1}[G' - (\alpha q\phi_k' + q'\phi_k)\phi_k^{\alpha-1}] \geqq \mu_k; \qquad V_{k-1}[G - q\phi_k^\alpha] \geqq \mu_k.$$

Thus G and G' have a "common factor" $\phi_k^{\alpha-1}$, with $\alpha - 1 > 0$. This factor is not a unit because ϕ_k is minimal in V_{k-1}. Thus Lemma 8.2 with §7, (1) gives

$$(5) \qquad V_{k-1}[R(G, G')] \geqq \mu_k \geqq \mu_{k-1} \qquad (k > 1).$$

For large k this is impossible. For if Γ_{k-1}, the cyclic value group of V_{k-1}, has the generator $\delta_{k-1} > 0$, while the group Γ_0 for V_0 is generated by $\delta_0 > 0$, then, because of §6, (3), and §5, (11),

$$(6) \qquad \delta_0/\delta_{k-1} = \exp V_{k-1} \leqq \deg \phi_k \leqq (\deg G)/(\text{proj } V_k).$$

Hence δ_{k-1} is bounded below by $\delta_0/\deg G$. But the sequence μ_i for $i \leqq k-1$ lies in Γ_{k-1} and is increasing (§7, (1)), so that it increases by steps of at least δ_{k-1}. Therefore $\mu_{k-1} \to \infty$ with k. But the field $K(\theta)$ was assumed separable, so that G has no multiple roots, whence $R(G, G') \neq 0$ and $V_{k-1}[R] = V_0 R$ is finite. Thus the inequality (5) is impossible for large k, and the assumption proj $V_k > 1$ is untenable for large k.

This proof can be used to estimate how soon proj V_k becomes 1.

If one combines (5) and (6) as indicated above, then

$$V_{k-1}[R(G, G')] \geqq [(k-2)\, \delta_0 \cdot \text{proj } V_k]/(\deg G).$$

This gives an upper bound for any k with proj $V_k > 1$. If we use the worst value, proj $V_k = 2$, in this bound and compute k' as the next larger integer, we find that *the integer k' of Theorem* 8.1 *may be taken as*

$$(7) \qquad k' = \left[\frac{\rho n}{2}\right] + 3,$$

where n is the degree of $G(x)$ and ρ the integer determined by $V_0[R(G, G')] = \rho\delta_0$.

Several improvements in this estimate are possible: (i), the term $\mu_k - \mu_{k-1}$, neglected in (5), can be estimated as not less than δ_0/n; (ii), if n is odd and proj $V_k = 2$, the last inequality of (6) can be improved, while the remaining cases of proj $V_k \geqq 3$ or n even, proj $V_k = 2$ can be treated by the original method. If this is carried out, one finds

$$(8) \qquad k' = \rho\left[\frac{n}{2}\right] + 2.$$

The whole argument can now be repeated with proj V_k replaced by the projection of the principal polygon for ϕ_k. This shows that once ϕ_k is chosen for $k \geqq k'$, the principal polygon has only one side, so that μ_k is completely determined. In other words, only the first half of the k'-th stage is needed for Theorem 8.1.

In the algebraic number case, ρ is the power to which the prime p under consideration divides the discriminant of G. If $\rho = 0$, then two stages suffice. This is essentially a part of the result of Dedekind, that under these conditions the prime ideal factors of p correspond to the irreducible factors $\phi_2(x)$ of $G(x)$ modulo[15] p. Presumably the estimate (8) could be improved by introducing the index (involving the non-essential discriminant divisors) of the original equation.

[15] R. Dedekind, *Ueber den Zusammenhang zwischen der Theorie der Ideale und der Theorie der höheren Kongruenzen*, Gesammelte Werke, vol. I, pp. 202–233.

9. **The degree of a value.** To interpret the relation (11) of §5 we need the notion of the "degree" of an absolute value. In an algebraic number field, the "inertial" degree of a prime ideal factor $\mathfrak{p}$ of a rational prime p is just the degree of the residue-class field of $\mathfrak{p}$ over the field of the integers mod p. To generalize to any value V of a ring S, use the ring of all "integers" $a \in S$ with $Va \geqq 0$, and call two integers a and b *congruent* mod V if $V(a - b) > 0$. The set of residue-classes of the integers with respect to this congruence forms as usual a ring, the *residue-class ring* S/V. If S is a field, so is S/V. If W is any extension of our original value V_0 to $K(\theta)$, the usual arguments show that the residue-class field $K(\theta)/W$ contains a subfield F_0 isomorphic to K/V_0 and that $K(\theta)/W$ is algebraic over this F_0. The *degree of* W is defined to be the degree, deg W, of $K(\theta)/W$ over F_0.

To compute the degree, we use the results of M, part II, which show that for a sequence of discrete inductive values $V_1, V_2, \cdots, V_k$ the residue-class ring of each V_i has the form of a polynomial ring $F_i[y]$, where F_i is an algebraic extension of $F_0 = K/V_0$. Furthermore (M, Theorem 12.1) $F_1 = F_0$, while, for $i \neq 0$, F_{i+1} is an algebraic extension of F_i of a degree which is exactly the degree of ϕ_{i+1} considered as a polynomial in $\phi_i^{\tau_i}$. In other words (M, Theorem 12.1),

$$\text{degree}\,(F_{i+1}:F_i) = \deg \phi_{i+1}/(\tau_i \cdot \deg \phi_i) \qquad (i = 1, \cdots, k-1).$$

These formulas, combined with the interpretation of τ_i in §6, (2), give

$$\text{(1)} \qquad \text{degree}\,(F_k:F_0) = \frac{\deg \phi_k}{\tau_1 \tau_2 \cdots \tau_{k-1}} = \frac{\deg \phi_k}{\exp\,(V_{k-1})}.$$

These results can be extended to non-finite inductive values thus[16]:

THEOREM 9.1. *For a non-finite value $V_k = [V_{k-1}, V_k\phi_k = \infty]$ the residue-class ring $K[x]/V_k$ is isomorphic to a field F_k, which is an algebraic extension of F_{k-1} of a degree determined as in* (1), *where $F_{k-1}[y]$ is the residue-class field of V_{k-1}.*

Proof. Exactly as in the proof M, Theorem 12.1, F_k is defined as the set of all residue-classes modulo V_k which contain a polynomial $f(x)$ with $V_{k-1}f \geqq 0$. But if a polynomial $g(x)$ in any residue-class is divided by ϕ_k, giving

$$g(x) = q(x)\phi_k + r(x),$$

then the term $q\phi_k$ has value ∞, so that g and r belong to the same residue-class, while $V_{k-1}r = V_k r \geqq 0$. Hence F_k includes all residue-classes and is the residue-class ring. Its degree is found as in M, Theorem 12.1.

THEOREM 9.2. *If W, an extension of V_0 to $K(\theta)$, corresponds as in Theorem* 2.1 *to an inductive value V_k with $V_kG(x) = \infty$, then*

$$\text{(2)} \qquad (\exp W)\cdot(\deg W) = \deg \phi_k.$$

[16] Theorem 9.1, as well as the last paragraph of §4, was revised July 15, 1936.

The correspondence of W to V_k yields an isomorphism between the residue-class rings $K(\theta)/W$ and $K[x]/V_k$. Hence by (1) and the definition of the degree of W,

$$\deg W = \text{degree } (F_k : F_0) = (\deg \phi_k)/\exp V_{k-1}.$$

But since any $V_k f$ is either $+\infty$ or some value from V_{k-1}, the value-groups of V_k and V_{k-1} are identical, and V_{k-1}, V_k, and W have the same exponent. Therefore (2) results.

A similar interpretation holds for a limit-value $V_\infty = \lim V_k$. We first prove as in M, Theorem 14.1, that, as soon as $\deg \phi_k = \deg \phi_{k+1} = \cdots$, we have $F_k = F_{k+1} = \cdots$, and that this constant F_k is the residue-class ring $K[x]/V_\infty$. As before, this F_k is then also the residue-class field of the corresponding value W of $K(\theta)$. Consequently, using (1) again, we get

THEOREM 9.3. *If W is an extension of V_0 to $K(\theta)$ which corresponds as in Theorem 2.1 to a limit-value $V_\infty = \lim V_k$ with $V_\infty G(x) = \infty$, then*

$$(\exp W)\cdot(\deg W) = \lim_{k\to\infty} \deg \phi_k, \tag{3}$$

and the limit on the right is actually attained for large k.

10. The totality of values. The existence theorem is

THEOREM 10.1. *There are only a finite number of extensions $W', W'', \cdots, W^{(s)}$ of a given discrete value V_0 of K to the separable field $K(\theta)$, where θ is a root of $G(x) = 0$. Furthermore,*

$$(\exp W')\cdot(\deg W') + \cdots + (\exp W^{(s)})\cdot(\deg W^{(s)}) = \deg G(x). \tag{1}$$

The relation (1) is a generalization of a well-known property of prime ideals. We first show that all W come from approximants. Every value W of $K(\theta)$ corresponds by Theorem 2.1 to a value of $K[x]$, which must be either an inductive value V_k or a limit-value V_∞. In the latter case, V_∞ is the limit of a sequence $V_1, V_2, \cdots$, in which each V_k is by Lemma 3.2 an approximant. In the former case, $V_k G(x) = \infty$ and V_{k-1} is by §2 and Lemma 3.2 a finite approximant. Since V_k is not finite, $V_k \phi_k = \mu_k = \infty$. Then only the multiples of ϕ_k have non-finite values, so that the last key ϕ_k must be $G(x)$ itself. This is the "terminating case" of Theorem 5.3. In this case there is only one sequence of approximants and hence only one value W of $K(\theta)$. The equation (2) of §9 thus gives the relation (1) above.

In the non-terminating case, we can construct one or more sequences of approximants $V_1, V_2, V_3, \cdots$. We must show that each such sequence gives a value W of $K(\theta)$. By Lemma 3.4

$$V_1 G(x) < V_2 G(x) < V_3 G(x) < \cdots, \tag{2}$$

while ultimately proj $V_k = 1$ and $\deg \phi_k$ is constant (Theorem 8.1 and §5, (5)). The index τ_k of each value-group Γ_{k-1} in the succeeding Γ_k is thus eventually unity (§6, (1) and (3)). Therefore all the values in (2) lie in some

one discrete group $\Gamma_{k'}$, so that $V_k G$ must approach ∞. The limit-value V_∞ then has $V_\infty G = \infty$, so that V_∞ corresponds to a value W of $K(\theta)$. The relation (1) for all these values follows from Theorems 5.2 and 9.3 because proj $V_k = 1$.

The complete limit-value V_∞ cannot be written down, but its essential properties can be calculated.

Theorem 10.2. *Each value $W^{(i)}$ of Theorem* 10.1 *is uniquely determined by an "approximant" inductive value $V_k^{(i)}$ of $K[x]$, for some $k = k'$. If it is possible to construct the irreducible factors of polynomials with coefficients in the residue-class field K/V_0, the approximants $V_k^{(i)}$ can be computed in a finite number of steps by finding certain slopes $\mu_j^{(i)}$ of the Newton polygons of $G(x)$ and certain key polynomials $\phi_j^{(i)}$ as the irreducible factors of $G(x)$ in various equivalence-decompositions. In this case one finds, in a finite number of steps,* (i) *the number s of extensions of V_0 to $K(\theta)$;* (ii) *the exponent and degree of each such $W^{(i)}$;* (iii) *the values $W^{(i)}\alpha$ for any previously given α in $K(\theta)$.*

This is a restatement of previous results, except for the last assertion, which gives a construction of the "prime ideal" decomposition of any α. If $\alpha = g(\theta) \neq 0$, then we need only compute $V_\infty g(x)$ for each limit value V_∞ involved. If for every k, $V_k g > V_{k-1} g$, the argument following (2) proves $V_\infty g = \infty$ and $\alpha = 0$. Otherwise $V_k g = V_{k-1} g$ for some k, so that V_k is not an approximant to $g(x)$ in the sense of Definition 3.3 and $V_\infty g = V_k g$ as in Lemma 3.2. Hence $W\alpha$ can be computed in k stages.

In the algebraic number case (K = the field of rationals) the construction of a prime ideal with inductive values can be extended to give a representation of the prime ideal as the greatest common divisor of integers. It can then also be proved that the "terminating case" of the construction arises whenever the prime p in question has only one prime ideal factor. The proof depends on the fact that every rational integer can be expressed as a sum of a finite number of terms cp^m, with $c = 0, 1, \cdots, p - 1$. Thence it can be argued that any approximant V_k with deg ϕ_k = deg G must ultimately lead to the terminating case.

It remains to connect our results with previous investigations on this topic. Ore[17] developed (Ore I) a construction for prime ideals in algebraic fields which for this special case is equivalent to the first $2\frac{1}{2}$ stages of our method, which involve the approximants V_2 and the key polynomials ϕ_3. This part of the construction does not suffice[18] for all equations $G(x)$. In a subsequent paper (Ore II, especially Kap. 2, §5) Ore made an extension equivalent to one more stage of our method, coupled with successive transformations of the defining equation $G(x)$, which have the effect of reducing several stages of our method

[17] Ore uses $\mu_1 = 0$, which is possible because θ is assumed integral.

[18] O. Ore, *Weitere Untersuchungen zur Theorie der algebraischen Körper*, Acta Math., vol. 45 (1925), pp. 145–160. Here it is proved that for every p and every algebraic field there "exists" a regular defining equation for which the second stage is sufficient. However, the existence proof is not constructive.

to one stage. This method is constructive and applies in all cases, but is justified only by appeal to another, more elaborate construction[19] of prime ideals in terms of congruences mod p^a. Berwick has developed[20] approximations equivalent to $2\frac{1}{2}$ stages of our method, and mentions the possibility of a third stage. The investigations of Wilson,[21] although they are formulated in terms of group-bases for ideals, are closely related to the first two stages of our method. However, if the method of successive approximations is to be universally applicable, it must be formulated in terms of an arbitrary number of steps; for, given an integer k and a prime p, an irreducible polynomial $G(x)$ can always be constructed so that the decomposition of p in the field defined by $G(x)$ will require more than k stages.

Our construction can also be employed to give a simple form to a number of irreducibility criteria,[22] to prove one of the fundamental theorems relating Hensel's p-adic numbers to prime ideals and to constructively establish the unique decomposition theorem in terms of the "Hauptordnungen" of Krull.[23] I plan to discuss some of these topics in a later paper.

HARVARD UNIVERSITY.

[19] O. Ore, *Ueber den Zusammenhang zwischen den definierenden Gleichungen und der Idealtheorie in algebraischen Körpern*, Math. Ann., vol. 96 (1926), pp. 313–352; vol. 97 (1927), pp. 569–598.

[20] W. E. H. Berwick, *Integral Bases*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 22.

[21] N. R. Wilson, *On finding ideals*, Annals of Math., vol. 30 (1928–29), pp. 411–428.

[22] S. MacLane, *Abstract absolute values which give new irreducibility criteria*, Proc. Nat. Acad. Sci., vol. 21 (1935), pp. 472–474; *The ideal-decomposition of rational primes in terms of absolute values*, Proc. Nat. Acad. Sci., vol. 21 (1935), pp. 663–667.

[23] W. Krull, *Idealtheorie*, p. 104.

ON THE CLOSURE OF $\{e^{i\lambda_n x}\}$

BY NORMAN LEVINSON

1. A set $\{e^{i\lambda_n x}\}$ is said to be closed $L^p(-\pi, \pi)$ if for any $f(x) \in L^p(-\pi, \pi)$

$$\int_{-\pi}^{\pi} f(x) e^{i\lambda_n x}\, dx = 0 \tag{1.0}$$

implies that $f(x)$ is equivalent to zero.

Here we will concern ourselves with the closure properties of the set

$$\{e^{i\lambda_n x}\} \qquad (-\infty < n < \infty),$$

where

$$\lim_{|n| \to \infty} \frac{n}{\lambda_n} = 1, \tag{1.1}$$

that is, the λ_n are positive for sufficiently large $n > 0$ and have density 1,[1] and a corresponding result holds for $n < 0$.

The question of the closure of such sets was first investigated by Wiener and Paley,[2] who considered closure in $L^2(-\pi, \pi)$ of even sets $(\lambda_{-n} = -\lambda_n)$. They made a special study of the set $\{1, e^{\pm i\lambda_n x}\}$, $n > 0$, where

$$|\lambda_n - n| \leqq B, \qquad n > 0, \tag{1.2}$$

and showed that if $B < \frac{1}{8}$, the set is closed $L^2(-\pi, \pi)$. Here it will be shown that for closure $L^2(-\pi, \pi)$ it suffices that $B \leqq \frac{1}{4}$.

First we shall obtain a general closure criterion. We shall use this criterion to get results under conditions of the type (1.2) and we shall show that these results are the best possible.

Our basic criterion is given by

THEOREM I. *Let $\{\lambda_n\}$ satisfy* (1.1). *Let $\Lambda(u)$ be the number of $|\lambda_n| \leqq u$. If*

$$\int_1^v \frac{\Lambda(u)}{u}\, du > 2v - \frac{p-1}{p} \log v - C \tag{1.3}$$

for some constant C, the set $\{e^{i\lambda_n x}\}$, $-\infty < n < \infty$, is closed $L^p(-\pi, \pi)$, $p \geqq 1$.

A corollary of Theorem I is

THEOREM II. *If*

$$|\lambda_n - n| \leqq \tfrac{1}{2} N + \frac{p-1}{2p} \qquad (-\infty < n < \infty), \tag{1.4}$$

Received June 4, 1936. The author is National Research Fellow.

[1] If the density is different from 1 (and not zero or infinite), the problem is reducible to this one by making a change of scale. In case densities do not exist, see Levinson, Proc. Camb. Phil. Soc., vol. 31 (1935), pp. 335–346.

[2] Wiener and Paley, *Fourier Transforms in the Complex Domain*, Am. Math. Soc. Coll. Pub., vol. XIX, Chap. VI.

the set $\{e^{i\lambda_n x}\}$, *if it is not closed, becomes closed on adjoining to it at most any* N *terms* $e^{i\alpha_n x}$, $1 \leqq n \leqq N$.

In particular, then, in order that a set be closed $L^p(-\pi, \pi)$, $p \geqq 1$, *it suffices that*

$$|\lambda_n - n| \leqq \frac{p-1}{2p}.$$

THEOREM III. *If we replace* (1.4) *by*

$$|\lambda_n - n| \leqq \tfrac{1}{2} N + \frac{p-1}{2p} + \delta, \tag{1.5}$$

where $\delta > 0$, *there exist sets* $\{e^{i\lambda_n x}\}$ *satisfying* (1.5) *which do not become closed when* N *terms are adjoined to them. Thus* (1.4) *is a best possible result.*

In connection with Theorems II and III the following result is of interest, although the proof is quite trivial. Note that it holds with no restrictions whatsoever on the $\{\lambda_n\}$.

THEOREM IV. *If the set* $\{e^{i\lambda_n x}\}$ *is closed* $L^p(-\pi, \pi)$, $p \geqq 1$, *it remains closed if we replace any* λ_n *by some other number.*

Obviously this result is equivalent to the one obtained by replacing "closed" by "unclosed" in the above statement, for if an unclosed set becomes closed, we apply Theorem IV to the closed set and obtain a contradiction at once.

2. **Proofs of the theorems** will now be given.

Proof of Theorem I. Let us suppose the theorem is not true. There exists an $f(x) \in L^p(-\pi, \pi)$ such that

$$H(w) = \int_{-\pi}^{\pi} f(x)\, e^{iwx}\, dx \tag{2.0}$$

has[3] zeros at $w = \lambda_n$, $-\infty < n < \infty$. Let us assume that none of the λ_n is zero. (How to proceed if one of them is zero will be obvious.)

Let

$$F(w) = \prod_{-\infty}^{\infty} \left(1 - \frac{w}{\lambda_n}\right) e^{w/\lambda_n}. \tag{2.1}$$

Since $H(w)$ vanishes at λ_n, $-\infty < n < \infty$,

$$\phi(w) = H(w)/F(w)$$

is an entire function. Denote the number of zeros of $H(w)$ not exceeding r in magnitude by $n(r)$, and those of $\phi(w)$ by $n_1(r)$. Then clearly

$$n_1(r) = n(r) - \Lambda(r). \tag{2.2}$$

It follows from Jensen's theorem that

$$\int_1^r \frac{n(u)}{u}\, du \leqq \frac{1}{2\pi} \int_0^{2\pi} \log^+ |H(re^{i\theta})|\, d\theta + A, \tag{2.3}$$

[3] The use of an entire function $H(w)$ in connection with the closure of trigonometric functions is due to Szász, Math. Annalen, vol. 77 (1916), pp. 482–496.

where A will be used throughout to represent various constants. By (2.0)

$$H(re^{i\theta}) = O(e^{\pi r |\sin\theta|}).$$

Using this in (2.3) we have

$$\int_1^r \frac{n(u)}{u}\,du \leqq 2r + A.$$

Using this and (1.3) in (2.2) we have

$$\int_1^r \frac{n_1(u)}{u}\,du \leqq \frac{p-1}{p}\log r + A.$$

Since $(p-1)/p < 1$, it is clear that $n_1(r) = 0$, or in other words that $\phi(w)$ has no zeros, that is, the zeros of $H(w)$ coincide with those of $F(w)$. By the Hadamard factorization theorem it follows therefore that $H(w) = ae^{bw}F(w)$, and in particular that

$$|H(iv)| = e^{cv}|aF(iv)|, \tag{2.4}$$

where a, b and c are constants.

From (2.0) we have for all sufficiently small $\epsilon > 0$,

$$|H(iv)| \leqq \left(\int_{-\pi}^{-\pi+\epsilon} + \int_{-\pi+\epsilon}^{\pi-\epsilon} + \int_{\pi-\epsilon}^{\pi}\right) e^{-vx}|f(x)|\,dt.$$

Using Hölder's inequality we have

$$\begin{aligned}
|H(iv)| &\leqq \left[\int_{-\pi+\epsilon}^{\pi-\epsilon} e^{-vxp/(p-1)}\,dx\right]^{(p-1)/p}\left[\int_{-\pi+\epsilon}^{\pi-\epsilon}|f(x)|^p\,dx\right]^{1/p} \\
&\quad + \left[2\int_{\pi-\epsilon}^{\pi} e^{x|v|p/(p-1)}\,dx\right]^{(p-1)/p}\left[\left(\int_{-\pi}^{-\pi+\epsilon} + \int_{\pi-\epsilon}^{\pi}\right)|f(x)|^p\,dx\right]^{1/p} \\
&\leqq e^{\pi|v|}|v|^{-(p-1)/p}\left\{2e^{-\epsilon|v|}\left[\int_{-\pi}^{\pi}|f(x)|^p\,dx\right]^{1/p}\right. \\
&\quad \left. + 2\left[\left(\int_{-\pi}^{-\pi+\epsilon} + \int_{\pi-\epsilon}^{\pi}\right)|f(x)|^p\,dx\right]^{1/p}\right\}.
\end{aligned}$$

For any $\delta > 0$ we can choose an ϵ so that

$$\left[\left(\int_{-\pi}^{-\pi+\epsilon} + \int_{\pi-\epsilon}^{\pi}\right)|f(x)|^p\,dx\right]^{1/p} = \delta.$$

Thus

$$|H(iv)| \leqq Ae^{\pi|v|}|v|^{-(p-1)/p}(e^{-\epsilon|v|} + \delta).$$

Or

$$\log|H(iv)| \leqq \pi|v| - \frac{p-1}{p}\log|v| + \log(e^{-\epsilon|v|} + \delta) + A.$$

From this it is clear, since we can choose δ arbitrarily small, that

$$\lim_{|v|\to\infty}\left(\log|H(iv)| - \pi|v| + \frac{p-1}{p}\log|v|\right) = -\infty.$$

Or using (2.4), we get

$$\lim_{|v|\to\infty}\left(\log|F(iv)| + cv - \pi|v| + \frac{p-1}{p}\log|v|\right) = -\infty. \tag{2.5}$$

On the other hand, by (2.1),

$$\log|F(iv)| = \tfrac{1}{2}\int_0^\infty \log\left(1+\frac{v^2}{u^2}\right)d\Lambda(u) = \int_0^\infty \frac{\Lambda(u)}{u}\,\frac{v^2}{u^2+v^2}\,du$$
$$= \int_0^\infty \frac{2v^2u}{(u^2+v^2)^2}\,du\int_0^u \frac{\Lambda(y)}{y}\,dy.$$

If we use (1.3), we have

$$\log|F(iv)| \geqq \int_0^\infty \frac{2v^2u}{(u^2+v^2)^2}\left(2u - \frac{p-1}{p}\log u - A\right)du$$
$$= \pi|v| - \frac{p-1}{p}\int_0^\infty \frac{2v^2u}{(u^2+v^2)^2}\log u\,du - A$$
$$= \pi|v| - \frac{p-1}{p}\log|v|\int_0^\infty \frac{2u}{(1+u^2)^2}\,du - A$$
$$= \pi|v| - \frac{p-1}{p}\log|v| - A.$$

But this contradicts (2.5).

Proof of Theorem II. Let us consider $\Lambda(u)$ for $u > N+1$. From (1.4) it follows that

$$\Lambda(u) \geqq 1 + 2\,[u - \tfrac{1}{2}N - (p-1)/2p], \qquad u > N+1,$$

or

$$\int_{N+1}^v \frac{\Lambda(u)}{u}\,du \geqq \int_{N+1}^v \frac{1+2[u-\frac{1}{2}N-(p-1)/2p]}{u}\,du$$
$$= 2\int_{N+1}^v \frac{u-\frac{1}{2}N-(p-1)/2p}{u}\,du$$
$$- 2\int_{N+1}^v \frac{u-\frac{1}{2}N-(p-1)/2p-[u-\frac{1}{2}N-(p-1)/2p]-\frac{1}{2}}{u}\,du.$$

Since $u - [u] - \frac{1}{2}$ is periodic, we get

$$\int_{N+1}^v \frac{\Lambda(u)}{u}\,du > 2\int_{N+1}^v \frac{u-\frac{1}{2}N-(p-1)/2p}{u}\,du - A$$
$$> 2v - N\log v - \frac{p-1}{p}\log v - A.$$

Now let us add N terms $e^{i\alpha_n x}$, $1 \leqq n \leqq N$, to $\{e^{i\lambda_n x}\}$ and denote the new set by $\{e^{i\mu_n x}\}$. Then clearly for sufficiently large u, $\mu(u) = \Lambda(u) + N$, where $\mu(u)$ is the number of $|\mu_n| \leqq u$. Thus

$$\int_1^v \frac{\mu(u)}{u}\, du > 2v - \frac{p-1}{p} \log v - A.$$

By Theorem I it follows at once that $\{e^{i\mu_n x}\}$ is closed $L^p(-\pi, \pi)$.

Before proving Theorem III we shall prove Theorem IV.

Proof of Theorem IV. Let $\{e^{i\lambda_n x}\}$ be closed and the set $\{e^{i\lambda_n x}\}$, $n \neq 0$, and $e^{i\alpha x}$, $\alpha \neq \lambda_0$, not be closed. There exists an $f(x) \in L^p(-\pi, \pi)$ not equivalent to zero such that

$$\int_{-\pi}^{\pi} f(x) e^{i\lambda_n x}\, dx = 0, \qquad n \neq 0,$$

$$\int_{-\pi}^{\pi} f(x) e^{i\alpha x}\, dx = 0.$$

Let us consider

$$g(x) = f(x) + i(\lambda_0 - \alpha) e^{-i\alpha x} \int_{-\pi}^{x} f(y) e^{i\alpha y}\, dy.$$

Clearly $g(x) \in L^p(-\pi, \pi)$. Moreover,

$$\int_{-\pi}^{\pi} g(x) e^{iux}\, dx = \int_{-\pi}^{\pi} f(x) e^{iux}\, dx + i(\lambda_0 - \alpha) \int_{-\pi}^{\pi} e^{i(u-\alpha)x} dx \int_{-\pi}^{x} f(y) e^{i\alpha y}\, dy,$$

or

$$\int_{-\pi}^{\pi} g(x) e^{iux}\, dx = \frac{u - \lambda_0}{u - \alpha} \int_{-\pi}^{\pi} f(x) e^{iux}\, dx. \tag{2.6}$$

It follows at once from this, on setting $u = \lambda_n$, that for any n

$$\int_{-\pi}^{\pi} g(x) e^{i\lambda_n x}\, dx = 0.$$

But $\{e^{i\lambda_n x}\}$ is closed $L^p(-\pi, \pi)$, and therefore $g(x)$ is equivalent to zero. But by (2.6) this means that

$$\int_{-\pi}^{\pi} f(x) e^{iux}\, dx = 0.$$

If we set $u = 0, \pm 1, \pm 2, \cdots$, this implies that $f(x)$ must also be equivalent to zero, contrary to our assumption.

Proof of Theorem III. First let us take the case when N is odd. We take

$$\begin{aligned} \lambda_n &= n + \tfrac{1}{2}N + \delta + (p-1)/2p \qquad && (n > 0), \\ \lambda_{-n} &= -\lambda_n && (n > 0), \\ \lambda_0 &= \tfrac{1}{2}N + \delta + (p-1)/2p. && \end{aligned} \tag{2.7}$$

By Theorem IV it does not of course matter where we take λ_0 (or any other finite number of λ_n). Let us set

$$\tfrac{1}{2} + \delta + (p - 1)/2p = t.$$

Then clearly $\cos^{2t-2} \tfrac{1}{2}x \in L^p(-\pi, \pi)$. Moreover, for $n \geqq 0$,

$$\begin{aligned}\int_{-\pi}^{\pi} & e^{i(n+t)x} \cos^{2t-2} \tfrac{1}{2}x \, dx \\ &= 2^{-2t+2} \int_{-\pi}^{\pi} e^{i(n+1)x}(1 + e^{ix})^{2t-2} \, dx \\ &= \lim_{r \to 1-0} 2^{-2t+2} \int_{-\pi}^{\pi} e^{i(n+1)x}(1 + re^{ix})^{2t-2} \, dx \\ &= \lim_{r \to 1-0} 2^{-2t+2} \sum_{k=0}^{\infty} r^k \binom{2t-2}{k} \int_{-\pi}^{\pi} e^{i(n+k+1)x} \, dx = 0.\end{aligned}$$

A similar result holds with $e^{-ix(n+t)}$, $n \geqq 0$. Thus $\cos^{2t-2} \tfrac{1}{2}x$ is orthogonal to $e^{\pm ix(n+t)}$, $n \geqq 0$. But the set $\{\pm(n + t)\}$, $n \geqq 0$, contains the set $\{\lambda_n\}$ defined in (2.7) and N additional terms. This proves Theorem III if N is odd.

If N is even, we proceed similarly but we now use $\sin \tfrac{1}{2}x \cos^{2t-1} \tfrac{1}{2}x$, where $t = \delta + (p - 1)/2p$. In this case $\{1, e^{\pm i(n+t)x}\}$, $n \geqq 1$, is orthogonal to $\sin \tfrac{1}{2}x \cos^{2t-1} \tfrac{1}{2}x$.

PRINCETON UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY.

REPRESENTATION OF POSITIVE HARMONIC FUNCTIONS

By Alfred J. Maria and Robert S. Martin

We are concerned with the problem of representing the positive harmonic functions in a given region, and are primarily interested here in pointing out the relevance to this problem of a number of other problems, some of which have been discussed in the literature. The representation, by means of the Poisson-Stieltjes integral, of the positive harmonic functions in a sphere is an instance of the type of representation with which we are concerned. The analytical technique customarily employed in establishing the Poisson-Stieltjes representation or one of its generalizations requires relatively stringent smoothness conditions (e.g., bounded curvature) upon the boundary of the region.[1] It is true that the criteria we here cite as sufficient for a solution of the representation problem are less explicitly connected with the nature of the boundary than are the usual conditions just referred to, and it does not seem a trivial problem to characterize intrinsically the regions for which these criteria are satisfied. Nevertheless, as we shall show elsewhere, our criteria are satisfied by classes of regions considerably broader than those to which the customary technique applies. This would seem to make it clear that the representation problem does not depend essentially on smoothness conditions, even in three or more dimensions where conformal mapping no longer serves as a *deus ex machina*.

In the present note we shall point out the criteria and give one two-dimensional application: a direct representation—that is, a representation not depending upon the intervention of conformal mapping—of the positive harmonic functions in a finitely multiply connected Jordan region. For simplicity we shall restrict the discussion to bounded regions and shall use two-dimensional language, but it is to be emphasized that, except in the application at the end, the argument is independent of the number of dimensions.

The representation in question is of the form

$$u(P) = \int_{A^*} f(S, P)\, d\mu(e_S), \tag{1}$$

where $u(P)$ is a non-negative harmonic function in a bounded region A, where A^* is the frontier of A, where $f(S, P)$ is a certain function which depends only upon the region A and which is defined for $S \in A^*$, $P \in A$, and where $\mu(e)$ is a finite, non-negative, completely additive function of Borel sets which vanishes

Received June 5, 1936.

[1] de la Vallée Poussin, *Propriétés des fonctions harmoniques dans un domaine ouvert limité par des surfaces à courbure bornée*, Annali della R. Scuola Normale Superiore di Pisa, (2), vol. 2 (1933), pp. 167–197; George A. Garrett, *Necessary and sufficient conditions for potentials of single and double layers*, Am. Jour. of Math., vol. 58 (1936), pp. 95–129.

in the complement of A^*. The integral is to be taken in the sense of Stieltjes-Radon.[2] The representation theorem is to the effect that for suitably chosen $f(S, P)$ equation (1) sets up a one-to-one correspondence between the non-negative harmonic $u(P)$ in A and the $\mu(e)$ which vanish in $-A^*$.

We shall denote by W the entire finite plane. If $M \subset W$, then $\bar{M}$, M^*, $-M$ will be respectively the closure, frontier, and complement of M. The symbols $u(P)$, $v(P)$, $\cdots$, etc., will *always* denote non-negative harmonic functions having as domain some bounded region (non-void, open, connected set).

We recall certain notions and results of which we shall make frequent use.[3]

For any bounded region A, let $m_A(e, P) = m(e, P)$ denote the mass distribution obtained by sweeping out a unit mass located at $P \epsilon A$ onto A^*. For a fixed $P \epsilon A$, $m(e, P)$ is a non-negative completely additive function of sets e, measurable Borel; the total mass is 1 and is located upon A^*. For a fixed e, $m(e, P)$ is a non-negative harmonic function of $P \epsilon A$. If every boundary point of A is regular, the solution of the continuous Dirichlet problem for boundary values $U(S)$ on A^* is given by

$$u(P) = \int_{A^*} U(S)dm(e_S, P). \tag{2}$$

The harmonic character in P of $m(e, P)$ shows that for any two fixed points P, P_0, $m(e, P)$ and $m(e, P_0)$ are as set functions each absolutely continuous with respect to the other, for they vanish on exactly the same sets e. In particular, this says that for any fixed P_0, a derivative function $[dm(e, P)]/[dm(e, P_0)]$ exists.[4] We shall eventually exhibit this derivative as a suitable choice for $f(S, P)$.

Suppose that the domain of $u(P)$ contains or is a region A. We shall denote by $E[u(P), A] = E(u, A)$ the set of all points $S \epsilon A^*$ for which $\overline{\lim}_{P \to S,\, P \epsilon A} u(P) > 0$. We may call $E(u, A)$ the *exceptional set* for $u(P)$ relative to A; it is the set of boundary points of A at which $u(P)$ does not take on continuously the boundary value 0. It is clear that if $E(u, A)$ is void, then $u(P)$ vanishes identically throughout A. Furthermore, if the boundary points of A are all regular, then for every Borel set e we have[5]

$$E[m(e, P), A] \subset \bar{e}. \tag{3}$$

The non-negative harmonic functions defined in a region A form a *normal family*; in particular, any collection of them which is bounded at some particular

[2] J. Radon, *Theorie und Anwendungen der absolut additiven Mengenfunktionen*, Wiener Sitzungsber., (1913), pp. 1295 ff.

[3] de la Vallée Poussin, *Extension de la méthode du balayage de Poincaré et problème de Dirichlet*, Annales de l'Institut H. Poincaré, vol. 2 (1932), pp. 169 ff. The results are more special than those needed here but the general result is valid and is in fact contained implicitly in results of N. Wiener, *Certain notions in potential theory*, Journal of Math. and Phys. (M.I.T.), vol. 3 (1924), pp. 24 ff.

[4] Radon, loc. cit., p. 1351.

[5] de la Vallée Poussin, loc. cit.

point of A—and thus at each point of A—is *compact* in the sense that every infinite subcollection contains a pointwise convergent sequence. More generally, if we speak of a family of subregions of A as *eventually covering* A whenever each closed subset of A is contained in all except possibly a finite number of these subregions, then again any infinite family of $u(P)$, whose domains eventually cover A and whose values are bounded at some point of A, contains a sequence pointwise convergent throughout A. In both cases the limit function is non-negative harmonic, and the convergence is uniform in every closed subregion of A.[6]

We shall say that a region A is approximated by a *nested family* of subregions $A_1, A_2, \cdots$, if (i) $\bar{A}_n \subset A_{n+1}$, (ii) $\lim_{n\to\infty} A_n = A$. It is clear that such a family eventually covers A.

Now let us turn to the representation problem. Consider the following conditions upon a bounded region A.

(α) The classical continuous Dirichlet problem is solvable for A; that is, every boundary point of A is regular.[7]

(β) Every boundary point of A admits the so-called 'principle of Picard'. That is to say, if $S_0 \epsilon A^*$ and $E(u, A) = E(v, A) = \{S_0\}$, then $u(P) = c\cdot v(P)$, where c is a positive constant.

(γ) For any closed set $B \subset A^*$, the condition $E(u, A) \subset B$ is a closed property of $u(P)$, that is to say, any limit element of a set of $u(P)$ having the property also has it.

Let us first observe that for regions A satisfying (α), the condition (γ) is equivalent to each of the following conditions:

(γ') For any closed $B \subset A^*$, the functions of a compact family of $u(P)$ with $E(u, A) \subset B$ are uniformly bounded near every $S \subset A^* - B$.

(γ'') For any closed $B \subset A^*$, the functions of a compact family of $u(P)$ with $E(U, A) \subset B$ take on uniformly the boundary value 0 at every $S \epsilon A^* - B$.

It is evident that (γ'') implies (γ). (α) and (γ') imply (γ''). For (α) says that a barrier[8] $V_S(P)$ exists at every $S \epsilon A^*$, and (γ') says that if $S \epsilon A^* - B$, we can majorize near S the $u(P)$ of any compact family of $u(P)$ for which $E(U, A) \subset B$ by a suitable positive multiple of $V_S(P)$. Finally, (γ) implies (γ'). If (γ') were not satisfied, there would be a convergent sequence of $u(P)$ $[E(u, A) \subset B]$ unbounded near some $S_0 \epsilon A^* - B$; thus a sequence of points $P_n \to S_0(P_n \epsilon A)$ and a convergent sequence of functions $u_n(P)$ $[E(u, A) \subset B]$ such that $u_n(P_n) > 2^n$. It can readily be verified that the functions $v_n(P) \equiv \sum_{m=1}^{n} 2^{-m} u_m(P)$ would form a compact, increasing, therefore convergent sequence, and that they together with their limit function would violate the

[6] A well known consequence of the Harnack inequality and the theorem of Ascoli. See for example P. Montel, *Leçons sur les Familles Normales de Fonctions Analytiques*, Paris, 1927, pp. 39 ff.

[7] O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929, p. 328.

[8] Kellogg, loc. cit., p. 327.

closure condition (γ). This completely establishes the equivalence of (γ), (γ'), (γ'') under the assumption of (α).

(ϵ) $f(S, P) = [dm(e, P)]/[dm(e, P_0)]$ can be defined for $S \in A^*$, $P \in A$ in such a way that it is continuous in S for fixed P and for fixed S is positive harmonic in P and equals unity at $P = P_0$.

(ζ) If $S_0 \in A^*$, then $\lim_{P \to S} f(S_0, P) = 0$ uniformly for all S outside a neighborhood of S_0.

THEOREM 1. *If A satisfies (α), (β) and (γ), then (ϵ) and (ζ) are satisfied.*

We separate the conclusion of the theorem into these two parts for technical reasons that will soon be clear.

In order to establish the theorem, let us consider a function $f(S, P)$ defined for $S \in A^*$, $P \in A$ which (i) is positive harmonic in P and 1 at $P = P_0$; (ii) satisfies $E[f(S, P), A] = \{S\}$. If such a function exists, it is by (β) clearly unique; by (i), (ii) and (γ'') it must satisfy (ζ); further, it must be continuous in S. To see the continuity, assume that $S_n \to S_0$ $(S_0, S_n \in A^*)$. Let $v(P)$ be any accumulation element of the sequence $\{f(S_n, P)\}$. For $n \geqq m$ we have $E[f(S_n, P), A] \subset B_m \equiv \sum_{k=m}^{\infty} \{S_k\} + \{S_0\}$. Application of (γ) yields $E(v, A) \subset B_m$ for all m. But $\prod_m B_m = \{S_0\}$. Therefore $v(P_0) = 1$ and $E(v, A) = \{S_0\}$. From (β) we get $f(S_0, P) = v(P)$. This says that for any sequence $S_n \to S_0$ there is a subsequence $S_n' \to S_0$ such that $f(S_n', P) \to f(S_0, P)$, which fact clearly implies continuity in S.

All we have, therefore, to show is that a derivative $[dm(e, P)]/[dm(e, P_0)]$ exists satisfying (i) and (ii) above.

It is convenient to compute the derivative over a net. Select in W rectangular cartesian coördinates x, y. Form the quadratic net consisting of all half-open squares $e^n_{p,q}$

$$\frac{p}{2^{n-1}} \leqq x < \frac{p+1}{2^{n-1}}, \qquad \frac{q}{2^{n-1}} \leqq y < \frac{q+1}{2^{n-1}},$$

where $p, q = 0, \pm 1, \pm 2, \cdots$; $n = 1, 2, \cdots$. For any $S \in W$ denote by $e_n(S)$ that $e^n_{p,q}$ which contains S. Form the sum e_0 of all $e^n_{p,q}$ for which $m(e^n_{p,q}, P_0) = 0$. Obviously $m(e_0, P_0) = 0$, and $-e_0 \subset A^*$.

Now $-e_0$ is dense in A^*. Otherwise e_0 would contain a subset e_1 relatively open in A^*. There would then be a continuous $V(S)$ defined in A^*, positive in e_1, and zero in $A^* - e_1$. The $v(P)$ determined from the boundary values $V(S)$ by means of (2) would vanish at P_0 and thus identically. This would contradict (α).

Now for any particular $S \in -e_0$ form the sequence of $g_n(S, P) \equiv [m(e_n(S), P)]/[m(e_n(S), P_0)]$. $g_n(S, P)$ is positive harmonic in P, 1 at $P = P_0$, and $E[g_n(S, P), A] \subset \overline{e_n(S)}$. Let $g(S, P)$ be any accumulation element of the $g_n(S, P)$. Applying (γ) and the fact that the $\overline{e_n(S)}$ form a descending sequence of sets, we get $E[g(S, P), A] \subset \overline{e_n(S)}$ for all n. Thus, since $\prod_n \overline{e_n(S)} = \{S\}$, $g(S, P)$

satisfies (i) and (ii) when those conditions are restricted to points $S \epsilon -e_0$. Furthermore, $g(S, P)$ is by its construction a derivative function[9] $[dm(e, P)]/[dm(e, P_0)]$, and this property will not be destroyed by an arbitrary extension of the domain of $g(S, P)$ to include points of the null set e_0. Extend $g(S, P)$ to all of A^* as follows. For any $S \epsilon A^*e_0$, choose a sequence $S_n \to S$, where the S_n lie in the dense set $-e_0$. Define $g(S, P)$ as some accumulation element of the sequence $\{g(S_n, P)\}$. By an almost exact reproduction of the argument at the first part of the proof we get $E[g(S, P), A] = \{S\}$. Thus the extended $g(S, P)$ satisfies (i) and (ii). The proof is therefore completed by taking $f(S, P) = g(S, P)$.

We may now deduce a number of immediate consequences of (ϵ) and (ζ).

THEOREM 2. *Suppose A satisfies (ϵ). If B is Borel and $\subset A^*$, and $\mu(e)$ is non-negative, completely additive, then*

$$u(P) = \int_B f(S, P)d\mu(e_S) \tag{4}$$

represents a non-negative harmonic function in A. Further, if A satisfies (ζ) and B is closed, then $E(u, A) \subset B$.

These statements follow from the approximation to the integral on the right of (4) by Riemann sums.

If A satisfies (ϵ) and $u(P)$ is representable by (1), where the $f(S, P)$ of (1) is that of Theorem 1, we call $u(P)$ *representable*.

THEOREM 3. *If A satisfies (ϵ), the totality of representable $u(P)$ form a closed class which contains every $u(P)$ taking on continuous boundary values over A^*.*

Suppose $u_n(P) \to u(P)$, where

$$u_n(P) = \int_{A^*} f(S, P)d\mu_n(e_S). \tag{5}$$

Since $f(S, P_0) = 1$, we get from (5) $u_n(P_0) = \int_{A^*} d\mu_n(e_S) = \mu_n(A^*)$. Thus the $\mu_n(e)$ have uniformly bounded total mass all contained in the compact set A^*. A subsequence $\{\mu_n'(e)\}$ therefore has a weak limit $\mu(e)$,[10] and we get

$$\begin{aligned} u(P) &= \lim_{n\to\infty} u_n(P) = \lim_{n\to\infty} \int_{A^*} f(S, P)d\mu_n(e_S) \\ &= \lim_{n\to\infty} \int_{A^*} f(S, P)d\mu_n'(e_S) = \int_{A^*} f(S, P)d\mu(e_S). \end{aligned} \tag{6}$$

[9] For a net such as we have used the extension of the Vitali covering theorem to completely additive set functions is readily established. From this one proceeds as with Lebesgue integrals.

[10] Radon, loc. cit., p. 1337.

This shows that the class of representable $u(P)$ is closed. If $u(P)$ takes on continuous boundary values $U(S)$, we have

$$\begin{aligned} u(P) &= \int_{A^*} U(S)dm(e_S, P) = \int_{A^*} U(S)d\int_{e_S} f(T, P)dm(e_T, P_0) \\ &= \int_{A^*} U(S)f(S, P)dm(e_S, P_0) = \int_{A^*} f(S, P)d\int_{e_S} U(T)dm(e_T, P_0) \qquad (7) \\ &= \int_{A^*} f(S, P)d\mu(e_S), \end{aligned}$$

where we have put $\mu(e) = \int_e U(T)dm(e_T, P_0)$.

Theorem 4. *If A satisfies (ϵ) and (ζ) and if $u(P)$ is representable, then the corresponding $\mu(e)$ is uniquely determined by $u(P)$.*

It is sufficient to prove that the value of $\mu(e)$ is determined for any closed subset e_0 of A^*. Let e_0 be such a set.

Let $A_1, A_2, \cdots$ be a nested family of approximating regions for A, which have, say, analytic boundaries. Or at least let them all satisfy condition (α). Such a family always exists.[11] We may assume that P_0 is in all A_n. Let $m_n(e, P) = m_{A_n}(e, P)$.

Define $M(P) = \operatorname{l.u.b.}_{S \epsilon e_0} f(S, P)$. For a positive ϵ define D_ϵ as the set of all $P \epsilon A$ for which $M(P) > \epsilon$. D_ϵ is an open set. (ζ) shows that D_ϵ cannot have a frontier point in $A^* - e_0$; i.e., that $\overline{D_\epsilon}A^* \subset e_0$.

Now for $S \epsilon A^*$ define

$$h_{n,\epsilon}(S) = \int_{A_n^* D_\epsilon} f(S, P)dm_n(e_P, P_0). \qquad (8)$$

We have

$$0 \leqq h_{n,\epsilon}(S) \leqq \int_{A^*} f(S, P)dm_n(e_P, P_0) = f(S, P_0) = 1. \qquad (9)$$

We now show that

$$\begin{aligned} &h_{n,\epsilon}(S) \geqq 1 - \epsilon && (S \epsilon e_0), \\ &\varlimsup_{n\to\infty} h_{n,\epsilon}(S) \leqq \epsilon && (S \epsilon A^* - e_0). \end{aligned} \qquad (10)$$

For $S \epsilon e_0$ and $P \epsilon A_n^* - A_n^* D_\epsilon$, we have $f(S, P) \leqq M(P) \leqq \epsilon$. Thus if $S \epsilon e_0$,

$$\begin{aligned} h_{n,\epsilon}(S) &= \left(\int_{A_n^*} - \int_{A_n^* - A_n^* D_\epsilon}\right) f(S, P)dm_n(e_P, P_0) \\ &\geqq 1 - \int_{A_n^* - A_n^* D_\epsilon} \epsilon\, dm_n(e_P, P_0) \geqq 1 - \epsilon, \end{aligned}$$

and thus the first of relations (10) is true.

[11] Kellogg, loc. cit., p. 319.

If $S \epsilon A^* - e_0$, the set $C_{S,\epsilon}$ of all $P \epsilon A$ for which $f(S, P) > \epsilon$ can have only one frontier point in A^*, namely, the point S. Thus $C_{S,\epsilon}$ has points in common with only a finite number of the sets $A_n^* D_\epsilon$; otherwise S would be a limit point of D_ϵ. Hence if $S \epsilon A^* - e_0$,

$$\begin{aligned}\overline{\lim_{n\to\infty}}\, h_{n,\epsilon}(S) &\leqq \overline{\lim_{n\to\infty}} \int_{A_n^* D_\epsilon(-C_{S,\epsilon})} f(S, P) dm_n(e_P, P_0) \\ &\leqq \int_{A_n^*} \epsilon\, dm_n(e_P, P_0) = \epsilon.\end{aligned}$$

Thus (10) is established.

Now form

$$T_{n,\epsilon} = \int_{A^*} h_{n,\epsilon}(S) d\mu(e_S). \tag{11}$$

We have

$$\begin{aligned}\overline{\lim_{n\to\infty}}\, T_{n,\epsilon} &\leqq \int_{A^*} \overline{\lim_{n\to\infty}}\, h_{n,\epsilon}(S) d\mu(e_S) \\ &= \left(\int_{e_0} + \int_{A^*-e_0}\right) \overline{\lim_{n\to\infty}}\, h_{n,\epsilon}(S) d\mu(e_S) \\ &\leqq \mu(e_0) + \epsilon\mu(A^* - e_0),\end{aligned} \tag{12}$$

and similarly

$$\underline{\lim_{n\to\infty}}\, T_{n,\epsilon} \geqq \int_{A^*} \underline{\lim_{n\to\infty}}\, h_{n,\epsilon}(S) d\mu(e_S) \geqq (1 - \epsilon)\mu(e_0). \tag{13}$$

(12) and (13) together show that

$$\mu(e_0) = \lim_{\epsilon\to 0} \left(\overline{\lim_{n\to\infty}}\, T_{n,\epsilon}\right). \tag{14}$$

Now, by Fubini's theorem

$$\begin{aligned}T_{n,\epsilon} &= \int_{A^*} h_{n,\epsilon}(S) d\mu(e_S) = \int_{A^*} \left[\int_{A_n^* D_\epsilon} f(S, P) dm_n(e_P, P_0)\right] d\mu(e_S) \\ &= \int_{A_n^* D_\epsilon} \left[\int_{A^*} f(S, P) d\mu(e_S)\right] dm_n(e_P, P_0) \\ &= \int_{A_n^* D_\epsilon} u(P) dm_n(e_P, P_0).\end{aligned} \tag{15}$$

The last expression does not explicitly involve $\mu(e)$. Therefore this, together with (14), show that the value of $\mu(e_0)$ is determined by $u(P)$.

So far we have shown that in a region A satisfying (ϵ) the $u(P)$ representable by (1) form a closed class which contains all $u(P)$ with continuous boundary values. If A also satisfies (ζ), the representation of a $u(P)$, if it exists, is unique. A fortiori these remarks are valid if A satisfies (α), (β), and (γ). For a com-

pleted representation theory we should need the representability of every $u(P)$. We shall not here investigate what minimum of conditions beyond (α), (β), (γ) are sufficient to secure this result, but rather shall add to them a fourth condition (δ) which we shall formulate presently. This procedure might appear somewhat unsatisfactory in view of the fact that, as will later turn out, the completed representation theorem under a very natural restriction implies (α) (β) and (γ), whereas there would at best be considerable difficulty in showing that it also implies the condition (δ). However, there is a very good technical justification for the condition (δ); namely, a sound technique, designed to establish (γ) for a class of regions, quite frequently yields (δ) as well, when suitably modified.

We now formulate the condition(δ).

(δ) There exists for A a nested family of approximating regions $A_1, A_2, \cdots$, each satisfying (α), (β), (γ), and their totality together with A fulfills the following condition: P_0 being a fixed point of A, and B being any closed set, the $u(P)$ which have as domain some A_n, which are less than some fixed bound at $P = P_0$, and which have their exceptional sets contained in B, are uniformly bounded near any point not in B.

It is clear that this condition states a kind of uniformity of the way in which the A_n satisfy the condition (γ) [in its equivalent form (γ')]. An application of the barrier condition analogous to that in the discussion of (γ), (γ') shows that when A satisfies (α) and (δ) and when there is given any convergent sequence of $u(P)$ whose domains are among the A_n, and eventually cover A, and whose exceptional sets are contained in B (closed), then the limit function of this sequence also has its exceptional set contained in B.

Now let A satisfy (α), (β), (γ), and let $A_1, A_2, \cdots$ be a nested family of approximating regions satisfying (α), (β), (γ). Select a fixed $P_0 \epsilon A_1$. Form for A relative to P_0 the function $f(S, P)$ of Theorem 1. Similarly form relative to P_0 for each A_n the corresponding function $f_n(S, P)$. Form the set $H \equiv A^* + \sum_{n=1}^{\infty} A_n^*$. H is closed. Define

$$F(S, P) = \begin{cases} f(S, P) & (S \epsilon A^*, P \epsilon A), \\ f_n(S, P) & (S \epsilon A_n^*, P \epsilon A_n). \end{cases} \tag{16}$$

THEOREM 5. *If A satisfies (α), (β), (γ) and (δ), then: (η) there exists for A a nested family of approximating regions satisfying (α), (β), (γ), such that the function $F(S, P)$ defined above is for fixed P continuous in $S \epsilon H$.*

To prove this, observe that as each A_n^* is at a positive distance from the rest of H, the only possible discontinuities (in S) of $F(S, P)$ would be at points of A^*, and these, if they existed, would be effective discontinuities only when approached over $H - A^*$. Thus if $F(S, P)$ were for some $P_1 \epsilon A$ discontinuous at $S_0 \epsilon A^*$, there would be a sequence $S_n \to S_0$ with $S_n \epsilon H - A^*$ and $\lim_{n\to\infty} F(S_n, P_1) \neq F(S_0, P_1)$. We could without loss of generality assume that $P_1 \epsilon A_1$

and that $S_n \epsilon A_n^*$. We could then let $v(P)$ be an accumulation element of the sequence $\{f_n(S_n, P)\}$. By an argument like that used in the proof of Theorem 1 we should have $v(P_0) = 1$, $E(v, A) = \{S_0\}$, and thus $v(P) = f(S_0, P)$. Hence

$$\lim_{n\to\infty} F(S_n, P_1) = \lim_{n\to\infty} f_n(S_n, P_1) = f(S_0, P_1) = F(S_0, P_1).$$

This would contradict $\lim_{n\to\infty} F(S_n, P_1) \neq F(S_0, P_1)$.

It is clear that (η) implies (ϵ) not only for A but for each of the approximating regions A_n.

THEOREM 6. *(η) implies the representability of every $u(P)$ in A.*

For a fixed $P \epsilon A_1$, $F(S, P)$ is continuous in S over the compact set H. Suppose $u(P)$ any non-negative harmonic function in A. Applying Theorem 3 to $u(P)$, relative to the region A_n, we get

$$u(P) = \int_{A_n^*} f_n(S, P)\, d\mu_n(e_S),$$

where the total mass of $\mu_n(e)$ is located upon A_n^*, and $\mu_n(A_n^*) = u(P_0)$. In particular, for $P \epsilon A_1$, we have

$$u(P) = \int_H F(S, P)\, d\mu_n(e_S).$$

The $\mu_n(e)$ have uniformly bounded total mass all contained in the compact set H. Hence a subsequence $\{\mu_{n_k}(e)\}$ has a weak limit $\mu(e)$. As each closed subset of A has points in common with only a finite number of A_n^*, the total mass of $\mu(e)$ must all be located upon A^*. Therefore, for $P \epsilon A_1$,

$$\begin{aligned} u(P) &= \lim_{k\to\infty} \int_H F(S, P)\, d\mu_{n_k}(e_S) \\ &= \int_H F(S, P)\, d\mu(e_S) = \int_{A^*} f(S, P)\, d\mu(e_S). \end{aligned}$$

Since, however, both sides of this equation represent harmonic functions throughout A, the equation must hold for all $P \epsilon A$.

We thus see that (η) implies the representability of every $u(P)$ by means of (1), where $f(S, P)$ is taken as $[dm(e, P)]/[dm(e, P_0)]$. (ϵ) and (ζ), à fortiori (η) and (ζ), guarantee that this representation is one-to-one. Under the assumption of (η) and (ζ) the representation has the further property, a consequence of Theorem 2, that a $u(P)$ takes on continuously the boundary value 0 at every point of A^* where the corresponding $\mu(e)$ has no mass. (A mass function is said to have no mass at a given point if the point has a neighborhood of zero mass.) We may call a representation having this last property *decomposable*. The condition for decomposability can readily be put in the form: if B is closed, and $\mu(e)$ is such that $\mu(eB) = \mu(e)$ identically in e, then for the corresponding $u(P)$ we have $E(u, A) \subset B$.

THEOREM 7. *If A is such that an $f(S, P) = [dm(e, P)]/[dm(e, P_0)]$ represents*

by (1) *in a one-to-one way all positive harmonic functions in* A, *and if this representation is decomposable, then* A *satisfies* (α), (β), *and* (γ).

Under these hypotheses $f(S, P)$ is for each $S \epsilon A^*$ a positive harmonic function of P. In fact, $f(S, P)$ is the function corresponding to a $\mu(e)$ due to a point mass located at S.

Furthermore, the one-to-one character of the representation implies a certain converse to the condition of decomposability; namely, if $u(P)$ is such that $E(u, A) \subset B$ (closed) and $\mu(e)$ is the corresponding mass function, then $\mu(eB) = \mu(e)$ identically in e. For assume B closed and $\subset A^*$. Let $u_0(P)$ be such that $E(u_0, A) \subset B$, and let $\mu_0(e)$ be the corresponding mass function. Let e_1 be any closed subset of $A^* - B$. Define $\mu_1(e) = \mu_0(ee_1)$, and let $u_1(P)$ correspond to $\mu_1(e)$. Clearly $\mu_1(ee_1) = \mu_1(e)$; hence by the decomposability $E(u_1, A) \subset e_1$. But since $\mu_1(e) \leqq \mu_0(e)$, we have $u_1(P) \leqq u_0(P)$, and therefore $E(u_1, A) \subset E(u_0, A) \subset B$. Thus $E(u_1, A) \subset Be_1$, which is void, and hence $u_1(P) \equiv 0$. Therefore $\mu_1(e) \equiv 0$. In particular, $\mu_0(e_1) = \mu_1(e_1) = 0$. As e_1 is any closed subset of $A^* - B$, $\mu_0(A^* - B) = 0$. Hence $\mu_0(e) = \mu_0(eA^*) = \mu_0[e(A^* - B)] + \mu_0(eB) = \mu_0(eB)$. Now turn to the condition (α). Let e_1 be relatively open in A^*. Put $e_2 = A^* - e_1$. Take $u_0(P) = m(e_2, P) = \int_{e_2} f(S, P)\, dm(e_S, P_0) = \int_{A^*} f(S, P) dm(e_S e_2, P_0) = \int_{A^*} f(S, P) d\mu_0(e_S)$. This $\mu_0(e)$ satisfies $\mu_0(e \cdot e_2) = \mu_0(e)$. Hence $E(u_0, A) \subset e_2$, and from this fact follows the solvability of the continuous Dirichlet problem.[12]

For (β), observe simply that if $E(u, A) = \{S_0\}$, then the corresponding $\mu(e)$ satisfies $\mu(e \cdot \{S_0\}) = \mu(e)$. In other words, this $\mu(e)$ is a point mass at S_0. But two such mass distributions for the same point S_0 are clearly multiples of one another, and the same must therefore be true of the corresponding harmonic functions.

Finally, for (γ), assume that $u_n(P) \to u(P)$ in A and $E(u_n, A) \subset B$ (closed). Let $\mu_n(e)$ correspond to $u_n(P)$. As $\mu_n(A^*) = u_n(P_0)$, the $\mu_n(e)$ have uniformly bounded total mass, and therefore a subsequence $\{\mu_{n_k}(e)\}$ has a weak limit $\mu(e)$. We therefore have

$$u(P) = \lim_{n \to \infty} u_n(P) = \lim_{k \to \infty} u_{n_k}(P) = \lim_{k \to \infty} \int_{A^*} f(S, P)\, d\mu_{n_k}(e_S)$$
$$= \int_{A^*} f(S, P)\, d\mu(e_S).$$

By what we proved above, $\mu_n(eB) = \mu_n(e)$. By the properties of a weak limit $\mu(eB) = \mu(e)$. Thus from the decomposability, $E(u, A) \subset B$. This concludes the proof.

As a corollary to this result, we see that (η) and (ζ) imply (α), (β), and (γ).

[12] de la Vallée Poussin, loc. cit., p. 203. The conclusions of the theorem in §44 are a special case of the statement above, and these conclusions are sufficient to give the solution of the continuous boundary value problem (loc. cit., p. 205, §45).

We now apply these results to finitely multiply connected Jordan regions. First, however, consider a special case. Consider a circular region C of center P_0 and radius R. The swept out mass $m(e, P)$ here has a continuous non-vanishing density with respect to arc length on C^*. Hence the derivative $f(S, P)$ is given by forming the quotient of these densities for the two functions $m(e, P)$, $m(e, P_0)$. It turns out that this quotient is $R^2 - PP_0^2/PS^2$, which is precisely the Poisson kernel. From this explicit expression it follows at once that C satisfies (ϵ) and (ζ). Furthermore, if we choose as approximating regions concentric circles interior to C, then (η) follows, again from the explicit expression. Thus there results a complete representation theory for the circle.[13]

We wish now to establish the conditions (α), (β), (γ), and (δ) for finitely multiply connected Jordan regions. This could be done quite readily by mapping such a region *conformally* upon a region whose boundaries are all circles. In these circumstances the conditions (α), (β), (γ), and (δ) would be invariant under the conformal mapping. For a region bounded by circles the swept out mass could be written down explicitly and use made of an argument similar to that above for the interior of a single circle. The approximating regions would be regions bounded by circles concentric with those of the original region. Since the approximating regions enter through (δ) but disappear in the final representation theory, it is irrelevant that they are restricted. However, it is of interest to observe that (δ) may be secured with much less stringent conditions on the approximating regions. It is for this reason that we choose a seemingly less obvious argument.

Let A be a finitely multiply connected Jordan region. Let $F_1, \cdots, F_k$ be the components of A^*.

It follows from known results that such an A satisfies (α).[14]

The above results show that (β) holds for a circular region, and a conformal mapping extends this to any simple Jordan region. A further extension to the present case may be argued as follows. Let S_0 be a boundary point of A, say a point of F_1. Let J be a simple Jordan arc joining two points of F_1 in A and separating S_0 from $F_2 + \cdots + F_k$ in A. There results a simple Jordan region A_0 having as boundary J plus a piece of F_1 containing S_0.

Now let $u(P)$ and $v(P)$ be two harmonic functions in A with $E(u, A) = E(v, A) = \{S_0\}$. Denote by $u_0(P)$ the harmonic function in A whose boundary values agree with $u(P)$ along J and are zero along $A_0^*F_1$. Similarly, define $v_0(P)$ in terms of $v(P)$. $u_0(P)$ and $v_0(P)$ are bounded, whereas $u(P)$ and $v(P)$ cannot be.[15] Therefore $E(u - u_0, A_0) = E(v - v_0, A_0) = \{S_0\}$. Now ($\beta$)

[13] Cf. G. Herglotz, *Über Potenzreihen mit positivem reellem Teil im Einheitskreis*, Ber. der Ges. Wiss. Leipzig, vol. 63 (1911), p. 501; G. C. Evans and H. E. Bray, *La formule de Poisson et le problème de Dirichlet*, Comptes Rendus, vol. 176 (1923), p. 1368; G. C. Evans, *Sur l'intégrale de Poisson*, Comptes Rendus, vol. 177 (1923), p. 241.

[14] For example, for simple Jordan regions it follows from a conformal mapping on a circle. The fact that regularity is a *local* property extends this to the present case.

[15] A harmonic function in A whose l.u.b. is $M < +\infty$ has a superior limit $> M - \epsilon$ at a set of positive capacity in A^*. Cf. Kellogg, loc. cit., p. 335.

for A_0 yields $u(P) - u_0(P) = c \cdot [v(P) - v_0(P)]$. Hence $u(P) - c \cdot v(P) = u_0(P) - c \cdot v_0(P)$. The right side of the last equation represents a bounded harmonic function in A_0, the left side a harmonic function which takes on zero boundary values, except possibly at S_0. Therefore in A_0, and hence throughout A, $u(P) - c \cdot v(P) = 0$.

As condition (γ) will follow from an obvious modification—in fact, simplification—of the method used for (δ), we establish (δ).

Let $P \to P' = \varphi(P)$ be any *topological* mapping of $\bar{A}$ upon a closed region whose boundary consists of a finite number of circles.[16] As a convention, if $M \subset \bar{A}$, then M' will be $\varphi(M)$; if $M' \subset \varphi(\bar{A})$, then M will be $\varphi^{-1}(M')$.

We show that (δ) holds when the approximating regions $A_1, A_2, \cdots$ are chosen as the (inverse) images of a nested family of regions $A_1', A_2', \cdots$ approximating A', their boundaries consisting of circles concentric with those of A'^*. We make such a choice of $A_1, A_2, \cdots$.

Let B be a closed subset of $\bar{A}$, S_0 a point of $A^* - A^*B$, P_0 a point of A_1. If (δ) were false, we could assume without loss of generality the existence of a sequence $\{u_n(P)\}$, with $u_n(P)$ defined in A_n, $u_n(P) = 1$, $E(u_n, A_n) \subset B$, and of a sequence $\{P_n\}$, with $P_n \in A_n$, $u_n(P_n) > n$, $P_n \to S_0$.

Now the set of points $P \in A_n$, where $u_n(P) > n$, will have a maximal connected open subset O_n containing P_n. For every $P \in A_n O_n^*$ we must have $u_n(P) = n$; otherwise O_n could be enlarged. Hence not all frontier points of O_n can lie in A_n, for we should then have $u_n(P) = n$ throughout O_n. Obviously no point of $A_n^* - A_n^* B$ could be a frontier point of O_n. Therefore $O_n^* B$ cannot be void. This shows that there must be a point $Q_n \in O_n$ distant $< n^{-1}$ from B, and a simple Jordan arc J_n joining P_n and Q_n in O_n. Along J_n we have $u_n(P) > n$. Thus no closed subset of A can have points in common with more than a finite number of J_n.

Suppose F_1 to be that component of A^* which contains S_0. We now say: there exists a simple Jordan arc K joining two points of F_1 in A in such a way that (i) K, together with a piece of F_1, bounds a simple Jordan region A_0; (ii) $\bar{A}_0(B + S_0) = 0$; (iii) there is a point $R \in A_0$ such that for infinitely many n it is true that a piece L_n of J_n divides A_0 into two simple Jordan regions, one of which, D_n, contains R and is included in A_n. The frontier of D_n thus consists of L_n and a piece of K.

The existence of such a K is readily established by carrying out the construction in the image region—taking there, for instance, K' (the eventual image of K) as a small arc of circle about a point of F_1'. In fact, if we take K_1' and K_2' as two sufficiently small arcs of circle with respective centers on either side of S_0' in a connected relatively open patch of F_1 free of B', then one or the other of K_1', K_2' can be used as K'.

[16] B. Kerékjártó, *Topologie*, Berlin, 1923, p. 121.

For infinitely many n we therefore have

$$(17)\qquad \begin{aligned} u_n(R) &= \int_{D_n^*} u_n(P)dm_{D_n}(e_P, R) \geqq \int_{L_n} u_n(P)dm_{D_n}(e_P, R) \\ &> \int_{L_n} n dm_{D_n}(e_P, R) = nm_{D_n}(L_n, R). \end{aligned}$$

But

$$(18)\qquad m_{D_n}(L_n, R) \geqq m_{A_0}(A_0^*F_1, R) > 0.$$

The last inequality is a consequence of the fact that $m_{A_0}(e, R)$ may be computed by first sweeping out a unit mass at R onto D_n^*, and then sweeping out the resulting mass on L_n onto A_0^*. All the mass on $A_0^*F_1$ is contributed by the second sweeping out. As total mass is conserved in sweeping out, the first part of the inequality follows. Furthermore, since $A_0^*F_1$ is of positive capacity, the second part of the inequality holds.

Combining (17) and (18), we see that the $u_n(P)$ are not bounded at the interior point R. This contradiction establishes (δ).

If in the above argument we choose the A_n all equal to A, and carry through the argument *mutatis mutandis*, (γ) follows.

Thus we have a completed representation theory for finitely multiply connected Jordan regions.[17]

INSTITUTE FOR ADVANCED STUDY.

[17] de la Vallée Poussin has proved a direct representation theorem for regions subject to the requirement that the irregular boundary points (i.e., points where the boundary curve is not of bounded curvature) constitute a perfect non-dense set in A^*. Cf. *Propriétés des fonctions harmoniques de deux variables dans une aire ouverte limitée par des lignes particulières*, Comptes Rendus, vol. 195 (1932), p. 92 ff.

FUCHSIAN GROUPS AND TRANSITIVE HOROCYCLES

By Gustav A. Hedlund

Let U denote the unit circle in the complex z-plane and let Ψ be its interior. The metric

$$ds^2 = \frac{4\,|\,dz\,|^2}{(1 - z\bar{z})^2}, \qquad z\bar{z} < 1 \tag{0.1}$$

defines a hyperbolic geometry in Ψ, the geodesics or *hyperbolic lines* of which are arcs of circles orthogonal to U. These hyperbolic lines will be designated as *H-lines*. The *hyperbolic distance* between two points of Ψ is defined as $\int ds$, where ds is given by (0.1) and the path of integration is the H-line segment joining these two points. The hyperbolic distance between P_1 and P_2 will be denoted by $H(P_1, P_2)$. The metric (0.1) is invariant under linear fractional transformations taking U into itself and Ψ into itself and these transformations transform hyperbolic lines into hyperbolic lines. Hence hyperbolic distance is invariant under all such transformations and these are the rigid motions of the geometry under consideration.

The curves in Ψ of constant geodesic curvature fall into four groups according to their geometrical properties (see e.g. Carathéodory,[1] pp. 22–25). If we denote geodesic curvature by g_c, these classes are as follows:

Class 1. $g_c = 0$. These are the H-lines and are arcs of circles orthogonal to U.

Class 2. $0 < g_c < 1$. These are the *hypercycles*. They are arcs of euclidean circles each of which meets U in two distinct points. The angle at which these curves meet U is uniquely determined by g_c and assumes all values between 0 and $\frac{1}{2}\pi$. The hypercycles are equidistant curves. That is, all the points of any given one are equidistant, in the hyperbolic sense, from the H-line which has the same end points on U.

Class 3. $g_c = 1$. These are the *horocycles* (oricycles). They are euclidean circles which are internally tangent to U.

Class 4. $g_c > 1$. These are the *hyperbolic circles* and lie entirely interior to U. All points of any given one are at the same H-distance from a fixed point in Ψ. These hyperbolic circles are also euclidean circles.

Let F be a fuchsian group with U as principal circle. If points congruent under F are considered identical, a two-dimensional manifold M, of constant negative curvature, is defined. The combinatorial topological properties of M are determined by F. If, in particular, F has a fundamental region lying, together with its boundary, in Ψ and F contains no elliptic transformations,

Received March 31, 1936.

[1] The references are to the bibliography at the end of the paper.

M is a closed orientable Riemannian manifold of genus greater than one and without singularities.

A curve of class C' on M is *transitive* if its elements are everywhere dense among the totality of elements on M. The question of the existence of transitive H-lines on M has been treated extensively. It is known (Koebe, p. 349) that there exist transitive H-lines if F is a fuchsian group of the first kind (Ford, p. 68). These are the groups which cease to be discontinuous at all points of U.

The H-lines are curves of constant geodesic curvature zero. What can be said with regard to the transitivity of curves of constant geodesic curvature not zero? This question is readily answered for the Classes 2 and 4 in the above classification. Class 4 is disposed of at once, for the curves in this class are closed, of finite length, and cannot be transitive. Any curve of Class 2, or hypercycle, is equidistant from a hyperbolic line and it is not difficult to show that a hypercycle is or is not transitive according as the H-line from which it is equidistant is or is not transitive.

The question of the transitivity or intransitivity of the curves of Class 3 or horocycles is not answered so readily. *It is the object of this paper to study the behavior of the horocycles, particularly with regard to transitivity.*

These results concerning the transitivity of the horocycles admit two applications. In the first place, it is possible to prove a mixture property of the flow defined by the hyperbolic lines on M. Secondly, there are derived properties concerned with the behavior of automorphic functions on circles internally tangent to U.

1. **The existence of transitive horocycles.** A horocycle is a euclidean circle which is internally tangent to U. It is completely determined by its euclidean radius and its point of contact with U. This point will be called the *point at infinity* on the given horocycle. The horocycle with euclidean radius r, $0 < r < 1$, and point at infinity Q will be denoted by $C(Q, r)$.

Two sets of points, within or on U, are *congruent* if there is a transformation of F taking one of these sets into the other. Either will be said to be a *copy* of the other.

Let $\bar{E}$ denote the set (x, y, φ), where $x^2 + y^2 < 1$ and $0 \leqq \varphi < 2\pi$. Any such point determines a point $P(x, y)$ of Ψ and a direction φ at this point, where directions at a point of Ψ are measured in the counterclockwise sense from the direction through the point parallel to the positive axis of reals. Conversely, a point in Ψ and a direction at this point determine a point in $\bar{E}$. A point in Ψ and a direction at this point will be called an *element*. Thus $\bar{E}$ is the space of elements. To define neighborhoods in $\bar{E}$, let $p_1(x_1, y_1, \varphi_1)$ be an arbitrary point of $\bar{E}$ and δ an arbitrary positive number. Let N_{p_1} be the set of points (x, y, φ) of $\bar{E}$ which satisfy the inequalities

$$H(P, P_1) < \delta, \qquad |\varphi - \varphi_1 + 2n\pi| < \delta,$$

for some integral n, where P is the point (x, y) and P_1 is the point (x_1, y_1). The set N_{p_1} defines a *neighborhood* of $p_1(x_1, y_1, \varphi_1)$. It is easily seen that $\bar{E}$, with neighborhoods thus defined, is a Hausdorff space.

Let $C(Q, r)$ be a directed horocycle. It has been defined as transitive if its elements are everywhere dense among the totality of elements on M, the two-dimensional manifold defined by identifying congruent points. This is equivalent to the following definition of transitivity.

Definition 1.1. *A directed horocycle $C(Q, r)$ is transitive if the totality of elements on $C(Q, r)$ and all its copies form a set which is everywhere dense in $\bar{E}$.*

Let P be any point other than Q of the horocycle $C(Q, r)$. The points P and Q divide $C(Q, r)$ into two parts, both of which will be termed *semihorocycles*. A point P of Ψ and a point Q of U determine two semihorocycles on both of which Q is the *point at infinity*. To distinguish between these, let C be a small circle with P as center and let A be the point in which C intersects the H-ray PQ. If C is traced out in the counterclockwise sense beginning at A, it intersects one of the semihorocycles determined by P and Q first. This one will be called a *right semihorocycle* and will be denoted by $SC_R(P, Q)$, the other a *left semihorocycle* and denoted by $SC_L(P, Q)$.

Definition 1.2. *A directed semihorocycle is transitive if the totality of elements on it and on all its copies forms a set which is everywhere dense in $\bar{E}$.*

It is evident that if a directed semihorocycle is transitive, the directed semihorocycle obtained by the removal of any finite segment of the given one is also transitive. If a directed semihorocycle is transitive, the directed horocycle of which it is a part is also transitive.

The remainder of this section is devoted to a proof of Theorem 1.1, which concerns the existence of transitive semihorocycles. A lemma which is of aid in the proof of this theorem is first derived. This lemma gives a criterion for determining when the transitivity of a set of semihorocycles implies the existence of an individual transitive semihorocycle. A set of semihorocycles is transitive if the totality of elements on all copies of all members of the set is everywhere dense in $\bar{E}$. This, of course, does not necessarily imply that any single member of the set is transitive.

Lemma 1.1. *Let P be a point of Ψ and Q_1Q_2 an interval of U. Let $Q_1'Q_2'$ be a subinterval of Q_1Q_2 and $SC_L[P, (Q_1'Q_2')]$ the set of directed left semihorocycles with P as initial point and with points at infinity in $Q_1'Q_2'$. If the set $SC_L[P, (Q_1'Q_2')]$ is transitive for every subinterval $Q_1'Q_2'$ of Q_1Q_2, there exists an infinite set of transitive left semihorocycles with P as initial point and with points at infinity everywhere dense in Q_1Q_2. The same result holds if left semihorocycles are replaced throughout by right semihorocycles.*

The neighborhoods N_{p_1} in $\bar{E}$ obtained by restricting the coördinates (x_1, y_1, φ_1) of p_1 to rational values and likewise for δ form a denumerable set of neighborhoods, $N_1, N_2, \cdots$. If a set of points in $\bar{E}$ has points in each member of this denumerable set of neighborhoods, it is evidently an everywhere dense set.

To prove the lemma, let $Q_1'Q_2'$ be an arbitrary subinterval of Q_1Q_2. Since

it is assumed that the set $SC_L[P, (Q_1'Q_2')]$ is transitive, there is in this set a directed semihorocycle with an element on it such that either this element or some copy of this element lies in N_1. Let Q'' be the point at infinity of this semihorocycle. Since the neighborhoods are open, there exists a closed interval $Q_1''Q_2''$ of U containing Q'' and such that *each* semihorocycle of the set

$$SC_L[P, (Q_1''Q_2'')]$$

either has on it an element of N_1 or an element with a copy in N_1. But the set $SC_L[P, (Q_1''Q_2'')]$ is transitive, so that the argument can be repeated with N_1 replaced by N_2. By repetition of this procedure, a sequence of closed intervals I_n, $n = 1, 2, \cdots$, of U is obtained with I_{i+1} contained in I_i, $i = 1, 2, \cdots$, and each of the directed semihorocycles of the set $SC_L[P, I_m]$ either has on it elements or there are copies of its elements which lie in the neighborhoods $N_1, N_2, \cdots, N_m$. This sequence of intervals has at least one point Q in common. But then the elements on the directed semihorocycle $SC_L[P, Q]$ and its copies form a set which has a point in each of the neighborhoods $N_1, N_2, \cdots$. This implies that this directed semihorocycle $SC_L[P, Q]$ is transitive. Since the interval $Q_1'Q_2'$ was an arbitrary subinterval of Q_1Q_2, the points Q with the property that $SC_L[P, Q]$ is transitive form an everywhere dense set in Q_1Q_2. This is the statement of the lemma.

If right semihorocycles are considered in place of left semihorocycles, the proof is similar.

There are fuchsian groups with U as principal circle for which none of the horocycles or semihorocycles is transitive. This is evidently true in the case of all fuchsian groups of the second kind (Ford, p. 68). These are the groups with limit points nowhere dense on U.

But in the case of fuchsian groups of the first kind, that is, groups with limit points everywhere dense on U, it is possible to prove the existence of transitive semihorocycles, and hence of transitive horocycles. In this case certain results are known which will be useful in the following. The transformations of any fuchsian group F with principal circle U are either hyperbolic with fixed points on U, parabolic with fixed point on U, or elliptic with fixed points inverse to U. If A and B are the fixed points on U of a hyperbolic transformation of F, the hyperbolic line AB is called the *axis* of the transformation. It will also be called a *periodic hyperbolic line*. It is known (Koebe, p. 349) that if F is of the first kind the periodic H-lines are everywhere dense among the totality of H-lines. This means that if I_1 and I_2 are arbitrary intervals of U, there is an axis of a transformation of F having one end point in I_1 and the other end point in I_2.

THEOREM 1.1. *If the group F is a fuchsian group of the first kind with principal circle U, P is an arbitrary point of Ψ and Q_1Q_2 is an arbitrary interval of U, there exist points C and D of Q_1Q_2 such that $SC_L(P, C)$ and $SC_R(P, D)$ are both transitive.*

If it can be shown that, Q_1Q_2 being an arbitrary interval of U, the sets $SC_L[P, (Q_1Q_2)]$ and $SC_R[P, (Q_1Q_2)]$ are both transitive, the theorem will follow

from Lemma 1.1. It is sufficient to prove that the set $SC_L[P, (Q_1Q_2)]$ is transitive. The proof is similar for the set $SC_R[P, (Q_1Q_2)]$.

Let h_{AB} be an arbitrary periodic directed H-line with A as initial point at infinity and B as terminal point at infinity. There exists a hyperbolic transformation of the group F with one of its fixed points interior to Q_1Q_2 and the other fixed point neither at A nor B. A properly chosen power of this transformation transforms h_{AB} into $h_{A'B'}$ with A' and B' both interior to Q_1Q_2.

Let R_0 be a point of $h_{A'B'}$ and s the hyperbolic arc length on $h_{A'B'}$, s being measured from R_0 and taken as positive in the positive sense of $h_{A'B'}$. Each point of $h_{A'B'}$ is then uniquely specified by a coördinate s, and R_s will denote the point determined by s. If s is sufficiently large in numerical value, R_s does not coincide with P and P and R_s determine a unique directed left semi-horocycle $SH_L(P, Q_s)$, with P as initial point, with R_s on it and with Q_s as point at infinity. As s becomes infinite in numerical value, Q_s approaches either A' or B', so that for s sufficiently large, Q_s lies in the interval Q_1Q_2.

It will be convenient to denote by $\{\varphi\}$, where φ is a real number, that unique number which satisfies the two conditions $\{\varphi\} \equiv \varphi$, mod 2π, and $0 \leqq \{\varphi\} < 2\pi$.

Let φ_s be the direction of $h_{A'B'}$ at R_s and φ'_s that of $SH_L(P, Q_s)$ at the same point. There are two possibilities with regard to the behavior of $\{\varphi'_s - \varphi_s\}$ as s becomes infinite, depending on the order of A' and B' on U. Either

Case I. $\lim_{s\to+\infty} \{\varphi'_s - \varphi_s\} = \frac{1}{2}\pi, \quad \lim_{s\to-\infty} \{\varphi'_s - \varphi_s\} = \frac{3}{2}\pi$, or

Case II. $\lim_{s\to+\infty} \{\varphi'_s - \varphi_s\} = \frac{3}{2}\pi, \quad \lim_{s\to-\infty} \{\varphi'_s - \varphi_s\} = \frac{1}{2}\pi$.

In either case, given δ, there exists an $\bar{s}$ such that $Q_{\bar{s}}$ lies in Q_1Q_2 and

$$| \{\varphi'_{\bar{s}} - \varphi_{\bar{s}}\} - \tfrac{1}{2}\pi | < \delta.$$

Let (x_a, y_a, φ_a) be an arbitrary element of $h_{A'B'}$. Since $h_{A'B'}$ is periodic, there exists an ω such that all the elements $(x_{a+m\omega}, y_{a+m\omega}, \varphi_{a+m\omega})$, $m = 0, \pm 1, \cdots$, are congruent, and hence copies of (x_a, y_a, φ_a). From the preceding, given $\delta > 0$, there exists an $\bar{m}$ such that $Q_{a+\bar{m}\omega}$ lies in Q_1Q_2 and $| \{\varphi'_{a+\bar{m}\omega} - \varphi_{a+\bar{m}\omega}\} - \frac{1}{2}\pi | < \delta$. This implies that for some integer n, $| \varphi'_{a+\bar{m}\omega} - \{\varphi_{a+\bar{m}\omega} + \frac{1}{2}\pi\} + 2n\pi | < \delta$. The elements $(x_{a+m\omega}, y_{a+m\omega}, \{\varphi_{a+m\omega} + \frac{1}{2}\pi\})$, $m = 0, \pm 1, \cdots$, are all copies of $(x_a, y_a, \{\varphi_a + \frac{1}{2}\pi\})$; hence given any element (x_a, y_a, φ_a) of h_{AB} and a neighborhood N of the element $(x_a, y_a, \{\varphi_a + \frac{1}{2}\pi\})$, there exists an element in N and either on one of the set $SC_L[P, (Q_1Q_2)]$ or on a copy of this set. But the element (x_a, y_a, φ_a) was an arbitrary element on a copy of an arbitrary periodic H-line and the preceding result can be stated as follows. Given an element (x, y, φ) of an arbitrary periodic hyperbolic line and an arbitrary neighborhood, N, of the element $(x, y, \{\varphi + \frac{1}{2}\pi\})$, there is an element on a copy of a member of the set $SC_L[P, (Q_1Q_2)]$ and in N. But the elements (x, y, φ) on the periodic H-lines are everywhere dense in $\bar{E}$ and the same must be true of the corresponding elements $(x, y, \{\varphi + \frac{1}{2}\pi\})$. Thus the elements on the copies of the set $SC_L[P, (Q_1Q_2)]$ must be everywhere dense in $\bar{E}$. The proof of Theorem 1.1 is complete.

The following theorem is an immediate consequence of Theorem 1.1.

THEOREM 1.2. *If F is a fuchsian group of the first kind, there exists an infinite set of transitive directed horocycles through any point of Ψ. The points at infinity on these transitive horocycles form an everywhere dense set on U.*

2. **The number of transitive horocycles.** The horocycle $C(Q, r)$ is determined by its point at infinity, Q, and its euclidean radius r. A directed horocycle will be called a *right horocycle*, $C_R(Q, r)$, if the sense of rotation on it is clockwise. If counterclockwise, it will be designated as a *left horocycle*, $C_L(Q, r)$. If a right horocycle is transitive, the left horocycle, which coincides with it except for sense, is transitive, and conversely.

The following theorem shows that the transitivity or intransitivity of a directed horocycle is determined by its point at infinity and is independent of its euclidean radius.

THEOREM 2.1. *If one directed horocycle with Q as point at infinity is transitive, all the directed horocycles with Q as point at infinity are transitive.*

It is sufficient to restrict the discussion to right horocycles and to show that the transitivity of $C_R(Q, r_1)$ implies that of $C_R(Q, r_2)$. Since the method of proof when $r_1 > r_2$ closely resembles that in the case $r_1 < r_2$, the proof will be given only with the assumption that $r_1 < r_2$.

The two horocycles $C_R(Q, r_1)$ and $C_R(Q, r_2)$ cut off equal hyperbolic lengths L_{12} on the set of H-lines with Q as point at infinity. Let $p_2(x_2, y_2, \varphi_2)$ be an arbitrary point of $\bar{E}$. There is a unique right horocycle $C_R(Q', r_2')$ having p_2 as an element of it. Let $C_R(Q', r_1')$, $r_1' < r_2'$, be the right horocycle such that $C(Q', r_1')$ and $C(Q', r_2')$ cut off equal hyperbolic lengths L_{12} on the hyperbolic lines with Q' as point at infinity. Let $p_1(x_1, y_1, \varphi_1)$ be the element of $C_R(Q', r_1')$ at the point $P_1(x_1, y_1)$ where it intersects the hyperbolic line determined by $P_2(x_2, y_2)$ and Q'. Since $C_R(Q, r_1)$ is assumed to be transitive, there exists a sequence of elements e_n^1, $n = 1, 2, \cdots$, which are copies of elements e_n, $n = 1, 2, \cdots$, of $C_R(Q, r_1)$ and are such that $\lim_{n\to\infty} e_n^1 = p_1(x_1, y_1, \varphi_1)$. If T_n denotes the transformation of F taking e_n into e_n^1, $n = 1, 2, \cdots$, the sequence $T_n[C_R(Q, r_1)] = C_R(Q_n, r_{1n})$, $n = 1, 2, \cdots$, evidently has the properties $\lim_{n\to\infty} Q_n = Q'$ and $\lim_{n\to\infty} r_{1n} = r_1'$. Under such conditions the right horocycle $C_R(Q', r_1')$ is said to be the *limiting right horocycle* of the sequence, and this is written $\lim_{n\to\infty} C_R(Q_n, r_{1n}) = C(Q', r_1')$. Since the transformations of F preserve hyperbolic distances, the sequence $T_n[C_R(Q, r_2)]$ must have $C_R(Q', r_2')$ as limiting right horocycle. This implies the existence of a sequence of copies of elements of $C_R(Q, r_2)$ having the element $p_2(x_2, y_2, \varphi_2)$ as limit element. But since p_2 was an arbitrary element in $\bar{E}$, $C_R(Q, r_2)$ is transitive.

Theorem 2.1 suggests a classification of the points of U. The point Q of U is *h-transitive* if all the directed horocycles with Q as point at infinity are transitive. The point Q of U is *h-intransitive* if none of the directed horocycles with Q

as point at infinity is transitive. From Theorem 2.1, all points of U are contained in these two categories.

As to the number of h-transitive points of U, so far it is only known that they form an infinite set. As a step in specifying precisely which points of U are h-transitive, the following theorem is derived.

THEOREM 2.2. *If F is a fuchsian group of the first kind, the end points of all axes of (hyperbolic) transformations of F are h-transitive.*

It is again sufficient to restrict the discussion entirely to right horocycles.

Let h_{AB} be the axis of a hyperbolic transformation T_H of F. If P is a finite point of h_{AB} and P' is the point $T_H(P)$, the hyperbolic distance ω between P and P' does not depend on how P is chosen on h_{AB}. Under the transformations T_H^n, $n = 0, \pm 1, \cdots$, the right horocycle $C_R(A, r_0)$ is transformed into a set $C_R(A, r_n)$, $n = 0, \pm 1, \cdots$, all with A as point at infinity. The pairs of right horocycles $C_R(A, r_n)$ and $C_R(A, r_{n+1})$, $n = 0, \pm 1, \cdots$, cut off equal hyperbolic lengths ω on all hyperbolic lines with A as point at infinity, hence, in particular, on the H-line through A and the origin. If a denotes the euclidean distance from the origin of a point at H-distance ω from the origin, there is a $C_R(A, r)$ in the set $C_R(A, r_n)$, $n = 0, \pm 1, \cdots$, such that $\frac{1}{2}(1 - a) \leqq r \leqq \frac{1}{2}$. If A' is a point of U which is a copy of A, the transformation of F taking A into A' transforms the set $C_R(A, r_n)$, $n = 0, \pm 1, \cdots$, into an infinite set of right horocycles with A' as point at infinity. Using the same argument on the new set, there is at least one member $C_R(A', r')$ of it such that $\frac{1}{2}(1 - a) \leqq r' \leqq \frac{1}{2}$, and $C_R(A', r')$ is a copy of $C_R(A, r_0)$.

From Theorem 1.2, there exists an h-transitive point Q of U. Since F is a fuchsian group of the first kind, there are hyperbolic transformations of F with fixed points arbitrarily close to Q, and hence there are copies of A arbitrarily close to Q. Let $A_0 = A, A_1, A_2, \cdots$ be a sequence of copies of A such that $\lim_{n\to\infty} A_n = Q$. If $C_R(A, r_0)$ is an arbitrary right horocycle with A as point at infinity, it has been shown that there exists a copy $C_R(A_n, r_n')$, $n = 0, 1, \cdots$, such that $\frac{1}{2}(1 - a) \leqq r' \leqq \frac{1}{2}$. The set of numbers r_n', $n = 0, 1, \cdots$, has at least one cluster value, $\bar{r}$, $\frac{1}{2}(1 - a) \leqq \bar{r} \leqq \frac{1}{2}$; hence there exists a subsequence $C_R(A_{n_i}, r_{n_i}')$, $i = 1, 2, \cdots$, of the set $C_R(A_n, r_n')$, $n = 0, 1, \cdots$, such that $\lim_{i\to\infty} A_{n_i} = Q$, and $\lim_{i\to\infty} r_{n_i}' = \bar{r}$. The sequence $C_R(A_{n_i}, r_{n_i}')$, $i = 1, 2, \cdots$, has the right horocycle $C_R(Q, \bar{r})$ as limiting right horocycle. The elements on the set $C_R(A_{n_i}, r_{n_i}')$, $i = 1, 2, \cdots$, have among their limit elements all elements of $C_R(Q, \bar{r})$. But the elements on $C_R(Q, \bar{r})$ and its copies are everywhere dense in $\bar{E}$, and hence the same must be true of the elements on the set $C_R(A_{n_i}, r_{n_i})$, $i = 1, 2, \cdots$, and on the copies of the members of this set. All such copies are copies of $C_R(A, r_0)$ and $C_R(A, r_0)$ must be transitive. This is the statement of Theorem 2.2.

Theorem 2.2 does not yield further information as to the number of h-transitive points of U. The end points of the axes form a denumerable set everywhere dense on U, but this set might coincide with the denumerable everywhere dense set of h-transitive points previously known to exist.

However, with the aid of Theorem 2.2 it is possible to give criteria for h-transitivity which immediately yield extensive results.

THEOREM 2.3. *If F is a fuchsian group of the first kind and there are copies of the horocycle $C(Q, r)$ with radii arbitrarily close to 1, Q is h-transitive.*

Again, the proof can be given with consideration of only right horocycles. From a sequence of copies of $C_R(Q, r)$ with radii approaching 1, a subsequence $C_R(Q_n, r_n)$, $n = 1, 2, \cdots$, can be chosen such that $\lim_{n\to\infty} Q_n = \bar{Q}$ and $\lim_{n\to\infty} r_n = 1$, where $\bar{Q}$ is a point of U. Let h_{AB} be an axis of a transformation of F such that $A \neq \bar{Q}$ and $B \neq \bar{Q}$. The fact that F is of the first kind implies the existence of such an axis. For all values of n sufficiently great, $C_R(Q_n, r_n)$ intersects h_{AB} in two points and the angle of intersection of $C_R(Q_n, r_n)$ and h_{AB} at either of these points approaches $\frac{1}{2}\pi$ as n becomes infinite. But if we use ω as previously defined, any point of h_{AB} is seen to have a copy in a fixed interval of h_{AB} of hyperbolic length ω. Hence there exists a sequence of copies of $C_R(Q, r)$, each of which intersects h_{AB} in a fixed interval and such that the angle of intersection approaches $\frac{1}{2}\pi$. This last implies that the points at infinity of the members of this sequence must either approach A or approach B. A subsequence $C_R(Q'_n, r'_n)$, $n = 1, 2, \cdots$, can be so chosen that $\lim_{n\to\infty} r'_n = r^* > 0$ and either $\lim_{n\to\infty} Q'_n = A$ or $\lim_{n\to\infty} Q'_n = B$. In either case the sequence has, from Theorem 2.2, a transitive right horocycle as limiting right horocycle and by the reasoning used in the proof of Theorem 2.2, $C_R(Q, r)$ must be transitive. This is identical with the statement that Q is h-transitive, thus proving Theorem 2.3.

THEOREM 2.4. *Let F be a fuchsian group of the first kind, Q a point of U and h_{OQ} the hyperbolic ray with the origin O as initial point and with Q as point at infinity. If there exists on h_{OQ} a sequence of points $O_0, O_1, \cdots$, such that $\lim_{n\to\infty} H(O, O_n) = +\infty$ and such that O_n has a copy O'_n, $n = 0, 1, \cdots$, with $H(O, O'_n)$ bounded, n arbitrary, Q is h-transitive.*

Consider the horocycle $C(Q, \frac{1}{2})$, which passes through O and contains h_{OQ}. Given $L > 0$, arbitrarily large, there exists a K such that the H-distance between O_K and any point of $C(Q, \frac{1}{2})$ exceeds L. Let T_K denote the transformation of F taking O_K into O'_K and let H' be an upper bound of the distances $H(O, O'_n)$. Assuming that L has been chosen greater than H', the H-distance from the origin to any point of the horocycle $T_K[C(Q, \frac{1}{2})]$ is not less than $L - H'$. But since L can be chosen arbitrarily large, this implies that there are copies of $C(Q, \frac{1}{2})$ with radii arbitrarily close to 1. From Theorem 2.3, Q is h-transitive.

With the aid of the criterion of Theorem 2.4, the h-transitivity of a large class of points of U is readily shown.

THEOREM 2.5. *If F has a fundamental region R_0 which, together with its boundary, lies entirely interior to U, all points of U are h-transitive.*

Under the hypothesis of the theorem, F must be of the first kind. This theorem is then an immediate consequence of Theorem 2.4, for any point of Ψ has a copy in R_0, hence at an H-distance from the origin less than a fixed con-

stant. Thus if F has a fundamental region which is bounded, in the hyperbolic sense, all horocycles are transitive.

What can happen if F is not so restricted, but is still of the first kind? It is easily seen that all points of U are no longer necessarily h-transitive. For F may contain parabolic transformations, and if Q is a fixed point of a parabolic transformation, and hence on U, all the horocycles with Q as point at infinity are periodic and cannot be transitive. The periodicity follows from the fact that a parabolic transformation with F as fixed point transforms each $C(Q, r)$ into itself.

But are these fixed points of parabolic transformations of F the only points of U which are not h-transitive? The answer can be shown to be in the affirmative in those cases where the fundamental region, R_0, has as boundary points on U only parabolic points.

THEOREM 2.6. *If F is of the first kind and if the only boundary points of R on U are parabolic points, all points of U, with the exception of those which are fixed points of parabolic transformations of F, are h-transitive.*

From the hypothesis of the theorem, F has a finite set of generators and R_0 a finite set of sides (Ford, p. 75) and thus the boundary points of R_0 on U must form a finite set $P_1, \cdots, P_m$. If the radii r_i, $i = 1, \cdots, m$, of the horocycles $C(P_i, r_i)$, $i = 1, \cdots, m$, are chosen sufficiently near 1, it is geometrically evident that any point of R_0 or its boundary and interior to U will be interior to some one of the set $C(P_i, r_i)$, $i = 1, 2, \cdots, m$. Denoting by C the set of horocycles consisting of $C(P_i, r_i)$, $i = 1, \cdots, m$, and all copies of these, any point in Ψ is interior to some member of the set C.

There exists a parabolic transformation T_i of F, with fixed point P_i, $i = 1, \cdots, m$. Hence each point of $C(P_i, r_i)$, with the exception of P_i, has a copy within H-distance D_i of the origin O, where D_i depends on $C(P_i, r_i)$ and not on the chosen point on it. Let D be a constant as great as any D_i, $i = 1, \cdots, m$. Any point of the set C which is not a point of U has a copy within H-distance D of the origin.

Now let Q be any point of U which does not belong to the set S_P consisting of $P_1, \cdots, P_m$ and the copies of these points. If Q' is any point other than Q of the H-ray OQ, it lies interior to one of the horocycles of the set C. But the ray $Q'Q$ cannot lie entirely in any one member of the set C, for this would imply that Q belonged to the set S_P. Hence $Q'Q$ must intersect one of this set of horocycles and has on it a point Q'' with a copy within H-distance D of the origin. From Theorem 2.4, the point Q is h-transitive.

If F is of the first kind, but has an infinite set of generators, Theorem 2.4 can be applied to prove the existence of at least a non-denumerable infinity of h-transitive points on U. For if Q is a point of U such that the geodesic rays with Q as point at infinity are transitive, the conditions of Theorem 2.4 are satisfied and Q is h-transitive. It is not difficult to show that a non-denumerable set of points of U have this property with regard to the geodesic rays. The precise analysis of h-transitivity in these cases requires further study, however.

3. **Asymptotic transitivity.** With the aid of the derived theorems concerning the transitivity of the horocycles, an interesting property of the flow defined by the hyperbolic lines on the manifold M can be shown to hold. To define the flow, we again consider the space $\bar{E}$ of elements (x, y, φ). Two such elements, (x, y, φ) and (x', y', φ'), are *congruent* if there is a transformation of F taking $P(x, y)$ into $P'(x', y')$ and the direction φ at P into the direction φ' at P'. Let E be the space obtained from $\bar{E}$ by considering congruent elements identical. Neighborhoods are defined in E by the definition of neighborhoods in $\bar{E}$ (Seifert-Threlfall, pp. 31–35). The space E is essentially the space of elements on M.

Let $A(X, Y, \Phi)$ be a point of E. This point is a set of elements and let (x, y, φ) be an arbitrary element of the set. The element (x, y, φ) defines a directed H-ray, r_h, namely, that one with (x, y) as initial point and with direction φ at that point. On r_h let (x_s, y_s) be the point at hyperbolic distance s from (x, y) and let φ_s be the direction of r_h at (x_s, y_s). The element (x_s, y_s, φ_s) and those congruent to it define a point $A_s(X_s, Y_s, \Phi_s)$ of E. The transformation or flow $A \to A_s$ of E into itself is evidently one-to-one and continuous, and depends continuously on the parameter s. If N denotes a point set of E, the transformation $A \to A_s$ transforms N into a set N_s.

Definition 3.1. *The flow $A \to A_s$ is asymptotically transitive (O) if, N and N^* being arbitrary open sets of E, there exists an S^* such that the set $N_s \cdot N^*$, $s > S^*$, is not empty.*

Theorem 3.1. *If F is of the first kind, the flow $A \to A_s$ is asymptotically transitive (O).*

It will be convenient to use the notation $\| \varphi \|$ to denote the member of the set $| \varphi + 2n\pi |$, $n = 0, \pm 1, \cdots$, which has the least numerical value.

Let $A(X, Y, \Phi)$ be an arbitrary point of N. This point determines an infinite set of points of $\bar{E}$ and let $(\bar{x}, \bar{y}, \bar{\varphi})$ be one of these. Since N is open, there exists a $\delta > 0$ such that all the elements $(\bar{x}, \bar{y}, \varphi)$, $\| \varphi - \bar{\varphi} \| < \delta$, determine points of N. The elements $(\bar{x}_s, \bar{y}_s, \varphi_s)$ are defined as before, φ being restricted to the set $\| \varphi - \bar{\varphi} \| < \delta$. The points $(\bar{x}_s, \bar{y}_s)$ thus defined form, for fixed s, an arc C_s of a hyperbolic circle with $(\bar{x}, \bar{y})$ as (non-euclidean) center. Let C_s be directed by directing the hyperbolic circle of which it is a part in the clockwise sense. The elements of C_s are given by the set $(\bar{x}_s, \bar{y}_s, |\varphi_s - \frac{1}{2}\pi|)$, $\| \varphi - \bar{\varphi} \| < \delta$.

Let $(X, Y, \Phi - \frac{1}{2}\pi)$ be the point of E defined by the set $(x, y, |\varphi - \frac{1}{2}\pi|)$, where (x, y, φ) is the set of $\bar{E}$ defining (X, Y, Φ). Let $\bar{N}^*$ be the open set of E obtained from the set N^* by replacing the points (X, Y, Φ) of N^* by the corresponding set $(X, Y, \Phi - \frac{1}{2}\pi)$. If it can be shown that there exists an S^* such that all C_s, $s > S^*$, determine sets of E with a point lying in $\bar{N}^*$, the desired theorem is proved. For then the elements $(\bar{x}_s, \bar{y}_s, \varphi_s)$, $s > S^*$, determine a set of E with a point lying in N^*.

To complete the proof, let Q_1 be the point at infinity on the H-ray determined by the element $(\bar{x}, \bar{y}, |\bar{\varphi} - \delta|)$ and let Q_2 be that on the H-ray determined by $(\bar{x}, \bar{y}, |\bar{\varphi} + \delta|)$. Let Q_1Q_2 be the interval of U consisting of the points at

infinity on the H-rays determined by the elements $(\bar{x}, \bar{y}, \varphi)$, $\|\varphi - \bar{\varphi}\| < \delta$. Since F is of the first kind, there exists an axis h_{BD} of a transformation T of F with B and D both interior to Q_1Q_2. Let T be such that it moves points away from B and towards D. Let P_1 be a finite point of h_{BD} and P_2 the point $T(P_1)$. Any finite point of h_{BD} has a copy in the interval P_1P_2, this copy being obtained by applying a properly chosen power of T.

For all s sufficiently large, C_s intersects h_{BD} in two points B'_s and D'_s, the notation being so chosen that B'_s lies between D'_s and B on h_{BD}. As s becomes infinite, the angle of intersection of C_s and h_{BD} at both B'_s and D'_s approaches $\frac{1}{2}\pi$.

If n is chosen properly, the transformation T^n transforms B'_s into a point B''_s in the interval P_1P_2 of h_{BD}. As s becomes infinite, the power of T required to transform B'_s into B''_s also becomes infinite and under these increasing powers of T the ends of C_s are transformed into points which approach D, in the euclidean sense. Thus, given $\epsilon > 0$, there exists an S^* such that for $s > S^*$ there is a copy, C''_s, of C_s intersecting the segment P_1P_2 of h_{BD} in B''_s at an angle which differs from $\frac{1}{2}\pi$ by less than ϵ and with the end points of this C''_s within euclidean distance ϵ of D. This copy is, of course, an arc of a euclidean, as well as a hyperbolic, circle.

Now consider the right horocycle $C_R(B'', D)$, B'' a point of the interval P_1P_2 of h_{BD}. By Theorem 2.2, $C_R(B'', D)$ is transitive, so that it has on it an element which determines a point of E in the set $\bar{N}^*$. Let C'' be an arc of a euclidean circle directed in the counterclockwise sense, C'' lying in Ψ, passing through B'' and having its end points near D. If ϵ is chosen sufficiently small and two conditions are fulfilled, namely, that the end points of C'' are within euclidean distance ϵ of D and the angle at which C'' intersects h_{BD} differs from $\frac{1}{2}\pi$ by less than ϵ, the arc C'' will have an element determining a point of E lying in $\bar{N}^*$. This is evident because C'' then approximates closely a large segment of $C_R(B'', D)$ and $\bar{N}^*$ is open. If C'' satisfies the same conditions, with the exception that it no longer necessarily intersects P_1P_2 in B'', but in a sufficiently small interval containing B'', it will still have an element determining a point of E lying in $\bar{N}^*$. By the Heine-Borel theorem, there exists an ϵ such that each such directed arc C'', with end points within distance ϵ of D, intersecting h_{BD} in any point of P_1P_2 and with angle of intersection with h_{BD} differing from $\frac{1}{2}\pi$ by less than ϵ, will have an element determining a point of E in $\bar{N}^*$.

But from the preceding, there exists an S^* such that every C_s, $s > S^*$, has a copy C''_s satisfying these conditions on C'', hence C_s, $s > S^*$, determines a set of points of E containing a point in $\bar{N}^*$. This completes the proof of Theorem 3.1.

4. **Double h-transitivity and automorphic functions.** Before taking into consideration the behavior of a function, automorphic with respect to F, on circular arcs internally tangent to U, it is desirable to extend the notion of h-transitivity.

DEFINITION 4.1. *The directed horocycle $C(Q, r)$ is doubly h-transitive if both*

of the directed semihorocycles $SC_R(P, Q)$ and $SC_L(P, Q)$ into which a point P of $C(Q, r)$ divides $C(Q, r)$ are both transitive.

THEOREM 4.1. *If one directed horocycle with Q as point at infinity is doubly h-transitive, all the directed horocycles with Q as point at infinity are doubly h-transitive.*

The proof of this theorem parallels that of Theorem 2.1 so closely that it does not seem necessary to give the details. It is a rather obvious consequence of the fact that two right (left) semihorocycles with the same point at infinity are, except possibly for a finite segment of one, equidistant curves, and if the elements of one determine a set which is everywhere dense in E, the same must be true of the other.

DEFINITION 4.2. *The point Q of U is doubly h-transitive if all the directed horocycles with Q as point at infinity are doubly h-transitive.*

This definition has significance because of Theorem 4.1.

Theorem 2.4 can be extended to double h-transitivity.

THEOREM 4.2. *Let F be a fuchsian group of the first kind, Q a point of U and h_{PQ} a hyperbolic ray with P, a point of Ψ, as initial point, and with Q as point at infinity. If there exists on h_{PQ} a sequence of points $P_1, P_2, \cdots$ such that $\lim_{n\to\infty} H(P, P_n) = +\infty$, and such that each of these points P_n has a copy P_n' with $H(O, P_n')$ bounded, n arbitrary, Q is doubly h-transitive.*

The proof will be given considering only directed left semihorocycles, since the proof for right semihorocycles is entirely similar. It will be sufficient to show that the directed left semihorocycle $SC_L(P, Q)$ with initial point P is transitive, for then every directed left semihorocycle with Q as point at infinity is transitive.

Since P_n has as copy P_n', the directed H-ray PQ has a copy passing through P_n'. Let e_n' be the element of this copy at the point P_n' and let T_n be the transformation of F taking P_n into P_n'. Since $H(O, P_n')$ is bounded, a subsequence e_n'', $n = 1, 2, \cdots$, of the sequence e_n', $n = 1, 2, \cdots$, can be chosen such that $\lim_{n\to\infty} e_n'' = e$, where e is an element of the set $\bar{E}$. Let Q_1 be the initial point at infinity of the directed hyperbolic line determined by e and let Q_2 be its terminal point at infinity. Then $\lim_{n\to\infty} T_n(P) = Q_1$ and $\lim_{n\to\infty} T_n(Q) = Q_2$, both in the euclidean sense.

Let $Q_1Q_2(\smile)$ be the segment of U traced out by a point starting at Q_1, tracing U in the counterclockwise sense, and terminating at Q_2. Since F is of the first kind, there is an axis h_{AB} of a transformation of F, with both A and B in $Q_1Q_2(\smile)$. For n sufficiently large, $T_n[SC_L(P, Q)]$ intersects h_{AB} in two points and as n becomes infinite, the angle of intersection at both of these points approaches $\frac{1}{2}\pi$. Denoting by $\bar{T}$ the transformation of F of which h_{AB} is an axis, let D be a point in Ψ and on h_{AB} and let $\bar{D} = \bar{T}(D)$. By applying a proper power of $\bar{T}$, there exists a copy of any curve in Ψ and intersecting h_{AB} such that the copy intersects h_{AB} at some point of the interval $D\bar{D}$. Hence, as we see by choosing n sufficiently large, there is a copy of $SC_L(P, Q)$ intersecting the

interval $D\bar{D}$ of h_{AB} at an angle nearly $\frac{1}{2}\pi$ and with the end points of this copy near B, in the euclidean sense. A sequence of these copies can be chosen such that the points of intersection of these copies with $D\bar{D}$ approach a point D' of $D\bar{D}$ as limit point, such that the angle of intersection approaches $\frac{1}{2}\pi$ and such that the end points of these copies approach B. But then there are copies of elements of $SC_L(P, Q)$ with any given element of the left horocycle $C_L(D', B)$ as limit element. Since $C_L(D', B)$ is transitive, by Theorem 2.2, $SC_L(P, Q)$ is also transitive.

By proofs entirely analogous to those of Theorems 2.5 and 2.6, the following two theorems are obtained.

THEOREM 4.3. *If F has a fundamental region R_0 which, together with its boundary, lies entirely interior to U, all points of U are doubly h-transitive.*

THEOREM 4.4. *If F is of the first kind and R_0 has only parabolic points on U, all points of U, with the exception of those which are fixed points of parabolic transformations of F, are doubly h-transitive.*

These theorems can be interpreted in terms of automorphic, or, in the case under consideration, fuchsian functions (Ford, p. 87).

THEOREM 4.5. *Let $F(z)$ be a fuchsian function with group F of the first kind. Let C be a circle internally tangent to U at P and $\widehat{PQ}$ any arc of C with Q as one end point. If F has a fundamental region R_0 which, together with its boundary, lies entirely interior to U, the values which $f(z)$ assumes on $\widehat{PQ}$ are everywhere dense among the totality of values assumed by $f(z)$ in Ψ. If F is of the first kind and R_0 has only parabolic points on U, the same property of $f(z)$ holds on any such arc $\widehat{PQ}$, provided Q is not a fixed point of a parabolic transformation.*

BIBLIOGRAPHY

1. Carathéodory, *Conformal Representation*, Cambridge Tracts in Mathematics and Mathematical Physics, 28.
2. Koebe, *Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen*, Fünfte Mitteilung, Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1930.
3. Ford, *Automorphic Functions*.
4. Seifert-Threlfall, *Lehrbuch der Topologie*.

BRYN MAWR COLLEGE.

STUDIES IN THE SUMMABILITY OF FOURIER SERIES BY NÖRLUND MEANS

BY MAX ASTRACHAN

I. Preliminary remarks and formulas

1. **Introduction.** In a paper published in 1932, E. Hille and J. D. Tamarkin [3][1] discussed the application of Nörlund means to the summation of Fourier series and certain associated series. They gave conditions for the effectiveness of the method in certain senses.

It is our purpose to consider further effectiveness problems of the theory for this method of summability. In particular, we shall consider its effectiveness for the summation of the Fourier series and conjugate Fourier series at points of "(C, α) continuity", and for the summation of the r-th derived series of both of these series. We shall also consider the strong summability of the two series when the partial sums are replaced by their Nörlund transforms.

2. **Nörlund means.** For a sequence $\{x_n\}$, the generalized Nörlund limit (if it exists) is defined as

$$(N, p_\nu)\text{-}\lim x_n = \lim_{n\to\infty} P_n^{-1}(p_n x_0 + p_{n-1}x_1 + \cdots + p_0 x_n), \tag{2.01}$$

where $\{p_\nu\}$ is a sequence of complex numbers such that $P_n \equiv p_0 + p_1 + \cdots + p_n \neq 0$. The conditions of regularity are

$$\sum_{k=0}^{n} |p_k| < C|P_n|, \qquad p_n/P_n \to 0, \tag{2.02}$$

where C is a fixed positive constant.

N. E. Nörlund [8] proved some properties of these means assuming $p_n > 0$ and $p_n/P_n \to 0$. Such a definition of limitation, however, had already been given by G. F. Woronoi [17], who assumed that $p_n > 0$ and that $n^{-\alpha}P_n$ is bounded for some value of α. We shall use the symbol (N, p_ν) to denote the Nörlund method of summation defined by the sequence $\{p_\nu\}$. If

$$p_n = \binom{\epsilon + n - 1}{n}, \qquad P_n = \binom{\epsilon + n}{n},$$

the corresponding method (N, p_ν) reduces to the Cesàro method (C, ϵ).

3. **Notation.** We shall consider functions $f(x)$ integrable in the sense of Lebesgue and periodic of period 2π. If

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt,$$

Received February 27, 1936.

[1] Numbers in square brackets refer to the bibliography at the end.

then $f(x)$ generates the Fourier-Lebesgue series

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{3.01}$$

whose conjugate series is

$$\sum_{n=1}^{\infty} (-b_n \cos nx + a_n \sin nx). \tag{3.02}$$

We designate by $s_n(x)$ and $\tilde{s}_n(x)$ the n-th partial sums of the series (3.01) and (3.02), respectively. The corresponding Nörlund transforms will be denoted by $N_n[f(x), p_\nu]$ and $\tilde{N}_n[f(x), p_\nu]$; and those of the r-th derivatives of the partial sums by $N_n^{(r)}[f(x), p_\nu]$ and $\tilde{N}_n^{(r)}[f(x), p_\nu]$. Further, we set

$$\varphi(t) \equiv f(x + t) + f(x - t) - 2f(x); \tag{3.03}$$

$$\varphi_\alpha(t) \equiv \frac{1}{\Gamma(\alpha)} \int_0^t (t - u)^{\alpha-1}\varphi(u)\, du \qquad (\alpha > 0); \tag{3.04}$$

$$\psi(t) \equiv f(x + t) - f(x - t); \tag{3.05}$$

$$\psi_\alpha(t) \equiv \frac{1}{\Gamma(\alpha)} \int_0^t (t - u)^{\alpha-1}\psi(u)\, du \qquad (\alpha > 0); \tag{3.06}$$

$$\tilde{f}(x) \equiv -\frac{1}{2\pi} \lim_{\epsilon \to 0} \int_\epsilon^\pi \psi(t) \cot \tfrac{1}{2}t\, dt; \tag{3.07}$$

$$\Phi_r(t) \equiv f(x + t) + (-1)^{r+1}f(x - t); \tag{3.08}$$

$$H_n(t) \equiv \frac{\cos \frac{1}{2}t - \cos (n + \frac{1}{2})t}{\sin \frac{1}{2}t}; \tag{3.09}$$

$$N_n(t) \equiv \frac{1}{2\pi P_n} \sum_{k=0}^{n} p_{n-k} \frac{\sin (k + \frac{1}{2})t}{\sin \frac{1}{2}t}; \tag{3.10}$$

$$\bar{N}_n(t) \equiv \frac{1}{2\pi P_n} \sum_{k=0}^{n} p_{n-k} H_k(t) \equiv \frac{1}{2\pi} \cot \tfrac{1}{2}t + \tilde{N}_n(t). \tag{3.11}$$

Finally, we shall denote by $N_n^{(r)}(t)$, $\bar{N}_n^{(r)}(t)$, and $\tilde{N}_n^{(r)}(t)$, the results obtained by differentiating $N_n(t)$, $\bar{N}_n(t)$, and $\tilde{N}_n(t)$, respectively, r times with respect to t.

4. **The Nörlund transforms.** It is well known that

$$s_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin (n + \frac{1}{2})(t - x)}{\sin \frac{1}{2}(t - x)}\, dt \tag{4.01}$$

and

$$\tilde{s}_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)H_n(x - t)\, dt. \tag{4.02}$$

Integrating (4.01) by parts and using (3.03), (3.04) with $\alpha = 1$, and (3.10), we have by (2.01) that

$$N_n[f(x), p_\nu] - f(x) = \int_0^\pi \varphi(t) N_n(t)\, dt = \frac{\varphi_1(\pi)}{2\pi P_n} \sum_{k=0}^{n} p_{n-k}(-1)^k$$

$$- \frac{1}{2\pi P_n} \int_0^\pi \frac{\varphi_1(t)}{\sin \frac{1}{2}t} \sum_{k=0}^{n} k p_{n-k} \cos (k + \tfrac{1}{2})t\, dt$$

$$- \frac{1}{4\pi P_n} \int_0^\pi \frac{\varphi_1(t)}{\sin \frac{1}{2}t} \sum_{k=0}^{n} p_{n-k} \cos (k + \tfrac{1}{2})t\, dt$$

$$+ \tfrac{1}{2} \int_0^\pi \varphi_1(t) \cot \tfrac{1}{2}t\, N_n(t)\, dt$$

$$\equiv L_1 + L_2 + L_3 + L_4. \tag{4.03}$$

Similarly, from (4.02) we have

$$\tilde{N}_n[f(x), p_\nu] - \tilde{f}(x) = \frac{-1}{2\pi P_n} \lim_{\epsilon \to 0} \frac{\psi_1(\epsilon)}{\sin \frac{1}{2}\epsilon} \sum_{k=0}^{n} p_{n-k} \cos (k + \tfrac{1}{2})\epsilon$$

$$+ \frac{1}{2\pi P_n} \int_0^\pi \frac{\psi_1(t)}{\sin \frac{1}{2}t} \sum_{k=0}^{n} k p_{n-k} \sin (k + \tfrac{1}{2})t\, dt$$

$$+ \frac{1}{4\pi P_n} \int_0^\pi \frac{\psi_1(t)}{\sin \frac{1}{2}t} \sum_{k=0}^{n} p_{n-k} \sin (k + \tfrac{1}{2})t\, dt$$

$$- \tfrac{1}{2} \lim_{\epsilon \to 0} \int_\epsilon^\pi \psi_1(t) \cot \tfrac{1}{2}t\, \tilde{N}_n(t)\, dt$$

$$\equiv \tilde{L}_1 + \tilde{L}_2 + \tilde{L}_3 + \tilde{L}_4. \tag{4.04}$$

Differentiating (4.01) r times with respect to x, we have by (2.01) with the notation of (3.10) that

$$\begin{aligned} N_n^{(r)}[f(x), p_\nu] &= \int_{-\pi}^{\pi} f(t) \frac{\partial^r}{\partial x^r} N_n(t - x)\, dt \\ &= (-1)^r \int_0^\pi [f(x + t) + (-1)^r f(x - t)] N_n^{(r)}(t)\, dt. \end{aligned} \tag{4.05}$$

Similarly, from (4.02),

$$\tilde{N}_n^{(r)}[f(x), p_\nu] = (-1)^{r+1} \int_0^\pi \Phi_r(t) \tilde{N}_n^{(r)}(t)\, dt. \tag{4.06}$$

5. **The** $(N, p_\nu)\cdot C_1$ **method.** This method of limitation is obtained by superimposing the method (N, p_ν) on the Cesàro means of order one. It is well known that if $\{\sigma_n(x)\}$ is the sequence of arithmetic means of the sequence $\{s_n(x)\}$, then

$$\sigma_n(x) - f(x) = \frac{1}{4(n + 1)\pi} \int_0^\pi \varphi(t) \left[\frac{\sin \frac{1}{2}(n + 1)t}{\sin \frac{1}{2}t}\right]^2 dt.$$

Integration by parts gives

$$\sigma_n(x) - f(x) = \frac{\varphi_1(\pi)}{4\pi(n+1)} \sin^2 (n+1)\frac{\pi}{2} - \frac{1}{8\pi}\int_0^\pi \frac{\varphi_1(t)}{\sin \frac{1}{2}t}\,\frac{\sin (n+\frac{1}{2})t}{\sin \frac{1}{2}t}\,dt$$

$$+ \frac{1}{4\pi(n+1)}\int_0^\pi \varphi_1(t) \cot \tfrac{1}{2}t \left[\frac{\sin (n+1)\frac{1}{2}t}{\sin \frac{1}{2}t}\right]^2 dt$$

$$\equiv C_1(n) + C_2(n) + C_3(n). \tag{5.01}$$

Hence if we denote by $N_n \cdot C_1[f(x), p_\nu]$ the Nörlund transform of $\sigma_n(x)$, it follows from (2.01) that

$$N_n \cdot C_1[f(x), p_\nu] - f(x) = \frac{1}{P_n}\sum_{k=0}^{n} p_{n-k}[C_1(k) + C_2(k) + C_3(k)]$$

$$\equiv C_1 + C_2 + C_3. \tag{5.02}$$

6. **Previous results.** For the convenience of the reader and for the sake of completeness, we state here certain results of Hille and Tamarkin. Their main theorem is the following:

THEOREM A. *A regular Nörlund method of summation (N, p_ν) is Fourier-effective if the generating sequence $\{p_\nu\}$ satisfies the following conditions:*

$$n\,|\,p_n\,| < C\,|\,P_n\,|, \tag{6.01}$$

$$\sum_{k=1}^{n} k\,|\,p_k - p_{k-1}\,| < C\,|\,P_n\,|, \tag{6.02}$$

$$\sum_{k=1}^{n} \frac{|\,P_k\,|}{k} < C\,|\,P_n\,|, \tag{6.03}$$

where C is a fixed positive constant independent of n.

They also prove the following

LEMMA 6.1. *If $\{p_\nu\}$ satisfies the condition*

$$\sum_{k=1}^{n} |\,p_k - p_{k-1}\,| = o(P_n), \tag{6.04}$$

then

$$\sum_{k=0}^{n} p_k e^{ikt} = o(P_n)$$

uniformly in t for $0 < \delta \leqq |\,t\,| \leqq \pi$.

It is also shown in their paper that (6.04) holds if (N, p_ν) satisfies, for example, conditions (6.02) and (6.03).

II. Summability at points of (C, α) continuity

7. **Definitions and results.** The Lebesgue sets of points associated with $\varphi(t)$ and $\psi(t)$ are those for which

$$\int_0^t |\,\varphi(u)\,|\,du = o(t), \qquad \int_0^t |\,\psi(u)\,|\,du = o(t),$$

and $f(x)$, $\bar{f}(x)$ exist and are finite, respectively. At such points the conditions of Theorem A are sufficient in order that the method (N, p_ν) sum the corresponding series to its correct value. We shall consider here summability at sets of points wider than the Lebesgue sets.

For $\alpha > 0$, the α-th integral of a Lebesgue integrable function $F(x)$ is defined as

$$F_\alpha(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-u)^{\alpha-1} F(u)\, du.$$

It is well known that $F_\alpha(t)$ exists for almost all t and is integrable; and that if $F_\alpha(t)$ exists at $t = t_0$, so does $F_\beta(t)$ for all $\beta > \alpha$. If $F_\alpha(t) = o(t^\alpha)$ as $t \to 0$, we say that $F(t)$ is continuous (C, α) at $t = 0$. It is well known that if $F(t)$ is continuous (C, α), it is also continuous (C, β) for all $\beta > \alpha$.

DEFINITION 7.1. A point x for which $f(x)$ has a definite value is said to be

(i) K_α regular if $\varphi(t)$ is (C, α)-continuous;

(ii) $\bar{K}_\alpha$ regular if $\psi(t)$ is (C, α)-continuous and $\bar{f}(x)$ exists and is finite.

DEFINITION 7.2. A summation method is said to be

(i) K_α effective if it sums the Fourier series of $f(x)$ to the correct value at all the K_α regular points;

(ii) $\bar{K}_\alpha$ effective if it sums the conjugate Fourier series of $f(x)$ to $\bar{f}(x)$ at the $\bar{K}_\alpha$ regular points.

We shall prove the following

THEOREM I. *A regular Nörlund method of summation (N, p_ν) is K_α and $\bar{K}_\alpha$ effective*[2] $(0 < \alpha \leqq 1)$, *if the generating sequence $\{p_\nu\}$ satisfies the following conditions:*

$$n\,|p_n| < C\,|P_n|, \tag{7.01}$$

$$\sum_{k=1}^{n} k\,|p_k - p_{k-1}| < C\,|P_n|, \tag{7.02}$$

$$\sum_{k=1}^{n} k(n-k)\,|p_k - 2p_{k-1} + p_{k-2}| < C\,|P_n|, \tag{7.03}$$

$$\sum_{k=1}^{n} \frac{|P_k|}{k^2} < C\,\frac{|P_n|}{n}, \tag{7.04}$$

where $p_{-1} = 0$ and C is a positive constant independent of n.

8. **Preliminary lemmas.** In this and the next two sections we shall assume that unless otherwise stated $f(x)$, $\bar{f}(x)$, and the generating sequence $\{p_\nu\}$ satisfy the conditions of Theorem I.

LEMMA 8.1. *If $\{p_\nu\}$ satisfies conditions* (7.03) *and* (7.04), *then*

$$\sum_{k=0}^{n} (n-k)\,|\Delta^2 p_k| = O(P_n/n) = o(P_n), \tag{8.01}$$

where $\Delta^2 p_k = \Delta p_k - \Delta p_{k-1} = p_k - 2p_{k-1} + p_{k-2}$, with $p_{-1} = p_{-2} = 0$.

[2] K_α effectiveness for the case $\alpha > 1$ will be discussed in §20.

Putting

$$U_n \equiv \sum_{k=1}^{n} k(n-k)\,|\Delta^2 p_k| = O(P_n), \qquad U_0 \equiv 0,$$

we have

$$\sum_{k=0}^{n} (n-k)\,|\Delta^2 p_k| = \sum_{k=1}^{n} \frac{1}{k}(U_k - U_{k-1}) = \sum_{k=1}^{n} \frac{U_k}{k(k+1)} + \frac{U_n}{n+1},$$

whence, by use of (7.04), the conclusion follows immediately.

LEMMA 8.2. *If* $\{p_\nu\}$ *satisfies conditions* (7.02) *and* (7.04), *then*

$$\sum_{k=1}^{n} |p_k - p_{k-1}| = O(P_n/n) = o(P_n). \tag{8.02}$$

The proof is similar to that of the previous lemma.

LEMMA 8.3. *If* $\{p_\nu\}$ *satisfies* (7.04), *then*

$$\sum_{k=1}^{n} \frac{|P_k|}{k} < C\,|P_n|.$$

This result follows at once from the hypothesis.

These lemmas show that a regular Nörlund method of summation (N, p_ν) which satisfies the conditions of Theorem I is automatically Fourier-effective. Therefore, since at the K_α and $\tilde{K}_\alpha$ regular points in the case $0 < \alpha \leqq 1$, we have $\varphi_1(t)\cot(\frac{1}{2}t) = o(1)$, $\psi_1(t)\cot(\frac{1}{2}t) = o(1)$, respectively, it follows that at all such points x, L_4 of (4.03) and $\tilde{L}_4$ of (4.04) are each $o(1)$ as $n \to \infty$. Further, by the well known Riemann-Lebesgue theorem and the regularity of the summation method, L_3 and $\tilde{L}_3$ are also $o(1)$.

LEMMA 8.4. *As* $n \to \infty$, $L_1 \to 0$ *and* $\tilde{L}_1 \to 0$.

From (4.03), by using Lemma 6.1, we get

$$L_1 = \frac{(-1)^n}{2\pi P_n}\varphi_1(\pi)\sum_{k=0}^{n} p_k e^{ikt}\Big|_{t=\pi} = o(1).$$

For $\tilde{L}_1$, we have from (4.04) that

$$|\tilde{L}_1| \leqq \lim_{\epsilon\to 0}\frac{|\psi_1(\epsilon)|}{2\sin\frac{1}{2}\epsilon}\,\frac{1}{\pi P_n}\sum_{k=0}^{n}|p_k| \leqq C\lim_{\epsilon\to 0}\frac{|\psi_1(\epsilon)|}{2\sin\frac{1}{2}\epsilon} = 0$$

at the $\tilde{K}_\alpha$ regular points.

The remaining integrals, L_2 and $\tilde{L}_2$, are essentially of the same type, for $\varphi_1(t)/\sin\frac{1}{2}t$ and $\psi_1(t)/\sin\frac{1}{2}t$ are functions of x and t which at the K_α and $\tilde{K}_\alpha$ regular points, respectively, approach zero with t; and the remaining parts of each integrand form the real and imaginary parts, respectively, of the complex-valued function

$$\mathfrak{M}_n(t) \equiv \frac{1}{2\pi P_n}e^{-i(n+\frac{1}{2})t}\sum_{k=0}^{n} p_k(n-k)e^{ikt}. \tag{8.03}$$

Our theorem will therefore be proved if we show that

$$\lim_{n\to\infty}\int_0^\pi g(t)\,\mathfrak{M}_n(t)\,dt = 0, \tag{8.04}$$

where $g(t) = o(1)$ as $t \to 0$.

9. **Estimates of the kernel.** We proceed to estimate the kernel (8.03).

LEMMA 9.1. *We have*

$$\mathfrak{M}_n(t) = O(n). \tag{9.01}$$

From the definition (8.03),

$$|\mathfrak{M}_n(t)| \leqq \frac{1}{2\pi|P_n|}\sum_{k=0}^{n}|p_k|(n-k) \leqq \frac{n}{2\pi|P_n|}\sum_{k=0}^{n}|p_k| = O(n),$$

in virtue of (2.02).

LEMMA 9.2. *If* $\{p_\nu\}$ *satisfies* (6.04), (8.01) *and* (8.02), *then*

$$\sum_{k=0}^{n} p_k(n-k)e^{ikt} = o(P_n) \tag{9.02}$$

uniformly in t *for* $0 < \delta \leqq |t| \leqq \pi$, δ *being fixed.*

Let us set $q_k = p_k(n-k)$ so that

$$\begin{gathered}\Delta q_k = q_k - q_{k-1} = (n-k)\,\Delta p_k - p_{k-1}\,,\\ \Delta^2 q_k = \Delta q_k - \Delta q_{k-1} = (n-k)\Delta^2 p_k - 2\Delta p_{k-1}, \quad q_n = 0, \quad \Delta q_n = -p_{n-1}\,.\end{gathered} \tag{9.03}$$

Then, putting

$$S_{k+1} = \sum_{\nu=0}^{k} e^{i\nu t} = \frac{1-e^{i(k+1)t}}{1-e^{it}}, \qquad S_1 = 1, \qquad S_0 = 0, \tag{9.04}$$

and applying Abel's partial summation method twice, we have

$$\begin{aligned}\sum_{k=0}^{n} q_k e^{ikt} &= \sum_{k=0}^{n}(q_{k-1}-q_k)S_k + q_n S_{n+1}\\ &= (1-e^{it})^{-1}\left[\sum_{k=0}^{n}(\Delta q_k)e^{ikt} - q_n e^{i(n+1)t}\right]\\ &= (1-e^{it})^{-2}\left[\sum_{k=0}^{n}(\Delta^2 q_k)e^{ikt} - (\Delta q_n)e^{i(n+1)t} - q_n(1-e^{it})e^{i(n+1)t}\right],\end{aligned}$$

where $q_{-1} = q_{-2} = 0$. Using (9.03), we get

$$\begin{aligned}\sum_{k=0}^{n} p_k(n-k)e^{ikt} = (1-e^{it})^{-2}\Bigg[&\sum_{k=0}^{n}(n-k)(\Delta^2 p_k)e^{ikt}\\ &- 2\sum_{k=0}^{n-1}(\Delta p_k)e^{i(k+1)t} + p_{n-1}e^{i(n+1)t}\Bigg].\end{aligned}$$

Hence under the assumptions of the lemma,

$$\left|\sum_{k=0}^{n} p_k(n-k)e^{ikt}\right| \leqq |1-e^{it}|^{-2}\left[\sum_{k=0}^{n}(n-k)\,|\,\Delta^2 p_k\,| + 2\sum_{k=0}^{n}|\,\Delta p_k\,| + |\,p_{n-1}\,|\right] = O(t^{-2})\cdot o(P_n),$$

where the "O" refers to $t \to 0$ and is independent of n, and the "o" refers to $n \to \infty$ and is independent of t. This completes the proof.

It remains to find suitable estimates for the kernel $\mathfrak{M}_n(t)$ in the interval (n^{-1}, δ). To do this we first consider the sum in (8.03).

Put

$$|\,p_n\,| = r_n, \qquad R_n = r_0 + r_1 + \cdots + r_n, \tag{9.05}$$

and introduce the step-functions

$$r(u) = r_{[u]}, \qquad R(u) = R_{[u]}, \tag{9.06}$$

where $[u]$ as usual denotes the largest integer $\leqq u$. Finally we put

$$\begin{aligned} &V_0 \equiv 0, \qquad V_n \equiv \sum_{k=1}^{n}|\,p_k - p_{k-1}\,|, \qquad V(u) \equiv V_{[u]},\\ &W_m \equiv \sum_{k=0}^{m}(n-k)\,|\,\Delta^2 p_k\,|, \qquad W(u) \equiv W_{[u]}. \end{aligned} \tag{9.07}$$

Consider

$$\sum_{k=0}^{n} p_k(n-k)e^{ikt} = \Sigma_1 + \Sigma_2 \qquad (t > 0). \tag{9.08}$$

where in Σ_1 k ranges over the integers $\leqq \tau = [1/t]$, and in Σ_2 over the integers $> \tau$ but $\leqq n$. Then

$$|\,\Sigma_1\,| \leqq \sum_{k=0}^{\tau}|\,p_\tau\,|(n-k) < n\sum_{k=0}^{\tau} r_k = nR_\tau = nR(t^{-1}). \tag{9.09}$$

Further, with our previous notation,

$$\begin{aligned}
\Sigma_2 &= -q_\tau S_{\tau+1} + \sum_{k=\tau+1}^{n}(q_{k-1}-q_k)S_k + q_n S_{n+1}\\
&= (1-e^{it})^{-1}\left[\sum_{k=\tau+1}^{n}(\Delta q_k)e^{ikt} + q_\tau e^{i(\tau+1)t} - q_n e^{i(n+1)t}\right]\\
&= \sum_{k=\tau+1}^{n}\frac{(\Delta^2 q_k)e^{ikt}}{(1-e^{it})^2} + \frac{(\Delta q_\tau)e^{i(\tau+1)t} - (\Delta q_n)e^{i(n+1)t}}{(1-e^{it})^2} + \frac{q_\tau e^{i(n+1)t} - q_n e^{i(n+1)t}}{1-e^{it}}\\
&= \frac{(n-\tau)p_\tau e^{i(\tau+1)t}}{1-e^{it}} + \frac{(n-\tau)(p_\tau - p_{\tau-1})e^{i(\tau+1)t} - p_{\tau-1}e^{i(\tau+1)t} + p_{n-1}e^{i(n+1)t}}{(1-e^{it})^2}\\
&\qquad + \sum_{k=\tau+1}^{n}\frac{(n-k)(\Delta^2 p_k)e^{ikt}}{(1-e^{it})^2} - 2\sum_{k=\tau+1}^{n}\frac{(\Delta p_{k-1})e^{ikt}}{(1-e^{it})^2}.
\end{aligned}$$

Hence if A denotes a generic constant independent of n and t,

$$|\Sigma_2| \leqq At^{-2}\Big[nt|p_\tau| + n|p_\tau - p_{\tau-1}| + |p_{\tau-1}| + |p_{n-1}| + \sum_{k=\tau+1}^{n} (n-k)|\Delta^2 p_k| + 2\sum_{k=\tau+1}^{n} |\Delta p_{k-1}|\Big];$$

and using (9.05), (9.06) and (9.07), we get

$$|\Sigma_2| \leqq At^{-2}[ntr(t^{-1}) + n|\Delta p_\tau| + r(t^{-1}-1) + r(n-1) + W(n) - W(t^{-1}) + 2V(n-1) - 2V(t^{-1}-1)].$$

Referring now to formulas (8.03), (9.08) and (9.09), together with the above, we have in virtue of (2.02) that

$$|\mathfrak{M}_n(t)| \leqq \frac{1}{2\pi|P_n|}\left|\sum_{k=0}^{n} p_k(n-k)e^{ikt}\right| \leqq A\sum_{j=1}^{7} M_{nj}(t),$$

where

$$M_{n1}(t) = \frac{n}{R(n)}R(t^{-1}), \qquad M_{n2}(t) = \frac{n}{tR(n)}r(t^{-1}), \qquad M_{n3}(t) = \frac{n}{t^2R(n)}|\Delta p_\tau|,$$

$$M_{n4}(t) = \frac{1}{t^2R(n)}r(t^{-1}-1), \qquad M_{n5}(t) = \frac{r(n-1)}{t^2R(n)},$$

$$M_{n6}(t) = \frac{1}{t^2R(n)}[W(n) - W(t^{-1})], \qquad M_{n7}(t) = \frac{1}{t^2R(n)}[V(n-1) - V(t^{-1}-1)].$$

10. **Proof of the theorem.** The truth of (8.04) under the conditions of the theorem now follows from the following lemmas.

LEMMA 10.1. *As $n \to \infty$, we have*

$$\int_0^{n^{-1}} g(t)\mathfrak{M}_n(t)\,dt = o(1).$$

By (9.01) and Definition 7.1, we have at once

$$\int_0^{n^{-1}} g(t)\mathfrak{M}_n(t)\,dt = O\left(\int_0^{n^{-1}} n\,dt\right) = o(1).$$

LEMMA 10.2. *For fixed $\delta > 0$, we have as $n \to \infty$ that*

$$\int_\delta^\pi g(t)\mathfrak{M}_n(t)\,dt = o(1).$$

This follows at once from Lemma 9.2 and formula (8.03).

In order to show finally that as $\delta \to 0$

$$\int_{n^{-1}}^{\delta} g(t)\mathfrak{M}_n(t)\,dt = o(1) \qquad (1/n < \delta)$$

at the points under consideration, it suffices to prove that

$$\int_{n^{-1}}^{\delta} |\mathfrak{M}_n(t)\,dt| < C. \tag{10.01}$$

That (10.01) is satisfied under the assumptions of the theorem follows from

LEMMA 10.3. *If $\{p_\nu\}$ satisfies the conditions of Theorem* I, *then*

$$\int_{n^{-1}}^{\delta} M_{nj}(t)\,dt \leqq M \qquad (j = 1, 2, \cdots, 7).$$

For $j = 1$, we have

$$\int_{n^{-1}}^{\delta} M_{n1}(t)\,dt = \frac{n}{R(n)}\int_{n^{-1}}^{\delta} R(t^{-1})\,dt = \frac{n}{R(n)}\int_{\delta^{-1}}^{n} \frac{R(s)}{s^2}\,ds.$$

This is bounded since $\dfrac{n}{R_n}\sum_{k=1}^{n}\dfrac{R_k}{k^2}$ is bounded by (7.04).

For $j = 2$,

$$\int_{n^{-1}}^{\delta} M_{n2}(t)\,dt = \frac{n}{R(n)}\int_{n^{-1}}^{\delta} r(t^{-1})\frac{dt}{t} = \frac{n}{R(n)}\int_{\delta^{-1}}^{n}\frac{r(s)}{s}\,ds.$$

This is bounded since $\dfrac{n}{R_n}\sum_{k=1}^{n}\dfrac{r_k}{k}$ is bounded by (7.04).

For $j = 3$,

$$\int_{n^{-1}}^{\delta} M_{n3}(t)\,dt = \frac{n}{R(n)}\int_{n^{-1}}^{\delta} |p_\tau - p_{\tau-1}|\frac{dt}{t^2} = \frac{n}{R(n)}\int_{\delta^{-1}}^{n} |p_{[s]} - p_{[s]-1}|\,ds$$

$$\leqq \frac{n}{R(n)}\sum_{k=0}^{n} |p_k - p_{k-1}|.$$

This is bounded by (8.02).

For $j = 4$, we have simply

$$\int_{n^{-1}}^{\delta} M_{n4}(t)\,dt = \frac{1}{R(n)}\int_{n^{-1}}^{\delta} r(t^{-1} - 1)\frac{dt}{t^2} = \frac{1}{R(n)}\int_{\delta^{-1}-1}^{n-1} r(s)\,ds < 1.$$

For $j = 5$, in virtue of (7.01),

$$\int_{n^{-1}}^{\delta} M_{n5}(t)\,dt = \frac{r(n-1)}{R(n)}\int_{n^{-1}}^{\delta}\frac{dt}{t^2} = O\left(\frac{nr_{n-1}}{R_{n-1}}\right) = O(1).$$

For $j = 6$,

$$\int_{n^{-1}}^{\delta} M_{n6}(t)\,dt = \frac{1}{R(n)}\int_{n^{-1}}^{\delta}[W(n) - W(t^{-1})]\frac{dt}{t^2} = \frac{1}{R(n)}\int_{\delta^{-1}}^{n}[W(n) - W(s)]\,ds$$

$$< \frac{1}{R(n)}\int_{0}^{n}[W(n) - W(s)]\,ds$$

$$= \frac{1}{R(n)}\int_{0}^{n} s\,dW(s) = \frac{1}{R_n}\sum_{k=1}^{n} k(n-k)\,|\Delta^2 p_k|.$$

This is bounded by (7.03).

Finally, for $j = 7$ as for $j = 6$, we have

$$\int_{n^{-1}}^{\delta} M_{n7}(t)\,dt = \frac{1}{R(n)} \int_{n^{-1}}^{\delta} [V(n-1) - V(t^{-1}-1)] \frac{dt}{t^2} < \frac{1}{R_n} \sum_{k=1}^{n} k \,|\, p_k - p_{k-1} \,|.$$

This is bounded by (7.02).

The proof of Theorem I is thus completed.

11. **The $(N, p_\nu)\cdot C_1$ theorem.** For this method of summation we have the following

THEOREM II. *The $(N, p_\nu)\cdot C_1$ method is K_α effective* $(0 < \alpha \leqq 1)$ *provided the generating sequence $\{p_\nu\}$ satisfies conditions* (6.01), (6.02) *and* (6.03).

From (5.01) it is obvious that $C_1(n) \to 0$. Further, the kernel of $C_3(n)$ is essentially a Fejér kernel, and at the points under consideration $\varphi_1(t) \cot \frac{1}{2}t = o(1)$, so that by a well known theorem, $C_3(n) \to 0$ as $n \to \infty$. Hence, since the generating sequence $\{p_\nu\}$ is regular, C_1 and C_3 of (5.02) approach zero.

Finally, since by hypothesis we also have $\varphi_1(t)/\sin\frac{1}{2}t = o(1)$ and the kernel of $C_2(n)$ is essentially a Dirichlet kernel, the result of Theorem A applies to show that $C_2 \to 0$ as $n \to \infty$.

A similar theorem can be proved for the $(N, p_\nu)\cdot C_1$ summability of the conjugate Fourier series.

III. The derived series

12. In a paper published in 1926 M. Jacob [4] proved that the r-th derived series of the Fourier series generated by a Lebesgue integrable function is summable by the Cesàro method $(C, r + \delta)$ to the value of the r-th generalized derivative of the function (in the sense of de la Vallée Poussin [5]) whenever this derivative exists. By means of an example he showed that the restriction $\delta > 0$ is necessary.

For the r-th derived conjugate series, summability in the case $r = 1$ has been considered by Plessner [10], Sayers [12], and Takahashi [15]. In the general case, Plessner [11] has defined an r-th derived conjugate function and stated without proof a theorem on the Cesàro summability of the series.

A. H. Smith [13] has introduced an r-th derived conjugate function to which he showed summability by the Bosanquet-Linfoot method of the r-th derived conjugate Fourier series. A. F. Moursund [6] has defined another such function and stated theorems for a method of summability introduced by F. Nevanlinna and extended by the former, as well as existence theorems for his function. He also stated such a theorem for the Bosanquet-Linfoot method [7].

We give here theorems concerning the summability of the r-th derived Fourier series and conjugate series for the Nörlund method (N, p_ν).

13. **Definitions and notation.** The generalized derivative in the sense of de la Vallée Poussin is defined as follows:

Definition 13.1. If at a point x the function $f(x)$ satisfies an equation of the form

$$\frac{f(x+t)+(-1)^r f(x-t)}{2} = \sum_{i=0}^{[\frac{1}{2}r]} f^{(r-2i)}(x)\,\frac{t^{r-2i}}{(r-2i)!} + \omega(x,t)\,\frac{t^r}{r!},$$

where $\omega(x, t) \to 0$ as $t \to 0$, then $f^{(r)}(x)$ is the r-th generalized derivative of $f(x)$.

From the existence of the r-th generalized derivative follows that of the $(r - 2j)$-th, $0 < 2j \leqq r$. Whenever the ordinary r-th derivative exists, it is identical with the generalized derivative $f^{(r)}(x)$.

With the notation of §3, we set

$$A_r(t) \equiv \Phi_r(t) - 2\sum_{i=0}^{[\frac{1}{2}(r-1)]} \frac{t^{r-1-2i}}{(r-1-2i)!}\,f^{(r-1-2i)}(x); \tag{13.01}$$

$$A_r^*(m) \equiv \int_t^m A_r(u)\,\frac{d^r}{du^r}\cot\tfrac{1}{2}u\,du; \tag{13.02}$$

$$B_r(n) \equiv 2(-1)^r \sum_{i=0}^{[\frac{1}{2}(r-1)]} f^{(r-1-2i)}(x)\int_0^\pi \frac{t^{r-1-2i}}{(r-1-2i)!}\,\frac{d^r}{dt^r}\tilde{N}_n(t)\,dt; \tag{13.03}$$

$$C_r \equiv \sum_{i=0}^{[\frac{1}{2}r]-1} f^{(r-1-2i)}(x)\sum_{j=i}^{[\frac{1}{2}r]-1}\frac{\pi^{2j-2i}}{(2j-2i+1)!}\,\frac{d^{2j+1}}{dt^{2j+1}}\cot\tfrac{1}{2}t\,\Big|_{t=\pi}; \tag{13.04}$$

$$D_r(n) \equiv \sum_{i=0}^{[\frac{1}{2}r]-1} f^{(r-1-2i)}(x)\sum_{j=i}^{[\frac{1}{2}r]-1}\frac{2\pi^{2j-2i+1}}{(2j-2i+1)!}\cdot \frac{d^{2j+1}}{dt^{2j+1}}\left[\frac{1}{2\pi}\cot\tfrac{1}{2}t - \tilde{N}_n(t)\right]_{t=\pi}; \tag{13.05}$$

$$U_i(r,n) \equiv (-1)^{r+1}\sum_{j=i}^{[\frac{1}{2}(r-1)]}\frac{\pi^{2j-2i+1}}{(2j-2i+1)!}\,N_n^{(2j)}(t)\,\Big|_{t=\pi}. \tag{13.06}$$

In any of the above expressions, whenever the upper limit of summation is less than the lower limit, that sum must be interpreted as zero. Thus $B_0(n) = 0$ and $C_r = D_r(n) = 0$ when $r = 0, 1$. We shall show later that $B_1(n) = 0$.

With this notation we define the r-th derived conjugate function $\tilde{f}^{(r)}(x)$ at a point where $f^{(r-1)}(x)$ exists by

Definition 13.2. For $r = 0, 1, 2, \cdots,$

$$\tilde{f}^{(r)}(x) \equiv \frac{(-1)^{r+1}}{2\pi}\lim_{\epsilon\to 0}\int_\epsilon^\pi A_r(t)\,\frac{d^r}{dt^r}\cot\tfrac{1}{2}t\,dt - C_r.$$

The function $\tilde{f}^{(0)}(x)$ is identical with $\tilde{f}(x)$ as given by (3.07), and $\tilde{f}^{(1)}(x)$ is essentially the same as $\tilde{f}'(x)$ as this limit is usually defined.

This definition was first given by Moursund. He proved[3] that if $\tilde{f}^{(r)}(x)$ exists, so does $\tilde{f}^{(r-2i)}(x)$, $0 < 2i \leqq r$, and that $\tilde{f}^{(r)}(x)$ exists (i) wherever $f^{(r+1)}(x)$ exists,

[3] In [6], p. 284; [7], p. 132. Note that we do not use quite the same expressions for the various symbols.

and (ii) almost everywhere when the $(r-1)$-th ordinary derivative is of bounded variation.

DEFINITION 13.3. A point x for which $f(x)$ has a definite value is said to be

(i) H_r regular if $f^{(r)}(x)$ exists and is finite;

(ii) $\tilde{H}_r$ regular if $\tilde{f}^{(r)}(x)$ exists and is finite.

DEFINITION 13.4. A summation method is said to be

(i) H_r effective if it sums the r-th derived Fourier series of $f(x)$ to $f^{(r)}(x)$ at all the H_r regular points;

(ii) $\tilde{H}_r$ effective if it sums the r-th derived conjugate Fourier series of $f(x)$ to $\tilde{f}^{(r)}(x)$ at all the $\tilde{H}_r$ regular points.

14. **Statement of results.** We set

$$\Delta^j p_k = \Delta^{j-1}p_k - \Delta^{j-1}p_{k-1}, \quad \cdots, \quad \Delta p_k = p_k - p_{k-1}, \quad \Delta^0 p_k = p_k, \tag{14.01}$$

where $p_{-k} = 0$, $k = 1, 2, \cdots$.

We shall prove the following two theorems.

THEOREM III. *A regular Nörlund method of summation (N, p_ν) is H_r effective if the generating sequence $\{p_\nu\}$ satisfies the following conditions:*

$$n^j |\Delta^{j-1}p_n| < C|P_n| \qquad (j = 1, 2, \cdots, r+1), \tag{14.02}$$

$$\sum_{k=1}^{n} k^{j-m-1}(n-k)^m |\Delta^{j-1}p_k| < C|P_n| \tag{14.03}$$

$$(j = 1, 2, \cdots, r+2;\ m = 0, 1, \cdots, r),$$

$$\sum_{k=1}^{n} \frac{|P_k|}{k^{r+1}} < C|P_n|n^{-r}, \tag{14.04}$$

where C is a fixed positive constant independent of n.

THEOREM IV. *A regular Nörlund method of summation (N, p_ν) is $\tilde{H}_r$ effective if the generating sequence $\{p_\nu\}$ satisfies conditions* (14.02), (14.03) *and* (14.04), *with r replaced by $r+1$.*

In what follows we shall assume that unless otherwise stated the conditions of Theorem III hold if we are dealing with the r-th derived Fourier series and those of Theorem IV hold if we are discussing the r-th derived conjugate series.

15. **Basic formulas.** In order to simplify the expressions (4.05) and (4.06) we shall use the following lemmas. The proofs depend on elementary computations and are omitted.

LEMMA 15.1.

$$N_n(t) = \frac{1}{2\pi P_n}\sum_{k=0}^{n} p_{n-k}[1 + 2(\cos t + \cos 2t + \cdots + \cos kt)];$$

$$\int_0^\pi N_n(t)\,dt = \tfrac{1}{2}; \qquad \frac{d^{2i+1}}{dt^{2i+1}} N_n(t) = 0 \ \textit{at}\ t = 0, \pi.$$

LEMMA 15.2. *For* $r - 2i \geqq 0$,

$$\int_0^\pi \frac{t^{r-2i}}{(r-2i)!} N_n^{(r)}(t)\,dt = \begin{cases} U_i(r, n) & (i > 0), \\ U_0(r, n) + \dfrac{(-1)^r}{2} & (i = 0). \end{cases}$$

LEMMA 15.3.

$$\bar{N}_n(t) = \frac{1}{\pi P_n} \sum_{k=0}^{n} p_{n-k}[\sin t + \sin 2t + \cdots + \sin kt];$$

$$\int_0^\pi \frac{d^{2j+1}}{dt^{2j+1}} \bar{N}_n(t)\,dt = 0; \qquad \frac{d^{2j}}{dt^{2j}} \bar{N}_n(t) = 0 \text{ at } t = 0, \pi.$$

It follows from this lemma and formula (13.03) that $B_1(n) = 0$.

LEMMA 15.4. $B_r(n) = C_r - D_r(n)$.

We proceed to transform the expression for $N_n^{(r)}[f(x), p_\nu]$. From (4.05), Definition 13.1 and Lemma 15.2, we have

$$\begin{aligned} N_n^{(r)}[f(x), p_\nu] &= 2(-1)^r \sum_{i=0}^{[\frac{1}{2}r]} f^{(r-2i)}(x) \int_0^\pi \frac{t^{r-2i}}{(r-2i)!} N_n^{(r)}(t)\,dt \\ &\quad + 2(-1)^r \int^\pi \omega(x, t) \frac{t^r}{r!} N_n^{(r)}(t)\,dt \\ &= f^{(r)}(x) + 2(-1)^r \sum_{i=0}^{[\frac{1}{2}r]} U_i(r, n) f^{(r-2i)}(x) \\ &\quad + 2(-1)^r \int_0^\pi \omega(x, t) \frac{t^r}{r!} N_n^{(r)}(t)\,dt. \end{aligned} \tag{15.01}$$

For $\bar{N}_n^{(r)}[f(x), p_\nu]$, we have from (4.06), using (13.01), (13.03), (13.04), (13.05) and (3.11), together with Lemmas 15.3 and 15.4, that

$$\begin{aligned} \bar{N}_n^{(r)}[f(x), p_\nu] &= (-1)^{r+1} \int_0^\pi A_r(t) \bar{N}_n^{(r)}(t)\,dt - C_r + D_r(n) \\ &= \frac{(-1)^{r+1}}{2\pi} \int_{n^{-1}}^\pi A_r(t) \frac{d^r}{dt^r} \cot \tfrac{1}{2}t\,dt - C_r + D_r(n) \\ &\quad + (-1)^{r+1} \int_0^{n^{-1}} A_r(t) \bar{N}^{(r)}(t)\,dt \\ &\quad + (-1)^{r+1} \left[\int_{n^{-1}}^{\delta} + \int_{\delta}^{\pi}\right] A_r(t) \bar{N}_n^{(r)}(t)\,dt \\ &\equiv \frac{(-1)^{r+1}}{2\pi} \int_{n^{-1}}^\pi A_r(t) \frac{d^r}{dt^r} \cot \tfrac{1}{2}t\,dt - C_r + D_r(n) + J_1 + J_2 + J_3. \end{aligned} \tag{15.02}$$

16. **Preliminary lemmas.** For the values of j indicated in (14.02) it follows at once that

$$| \Delta^{j-1} p_n | = o(P_n). \tag{16.01}$$

Further, from (14.04),

$$(16.02)\qquad \frac{n^\beta}{|P_n|}\sum_{k=1}^{n}\frac{|P_k|}{k^{\beta+1}} = O(1) \qquad (\beta = 0, 1, 2, \cdots, r).$$

Finally,

$$(16.03)\qquad \sum_{k=0}^{n}(n-k)^m|\Delta^{m+1}p_k| = O(P_n/n) = o(P_n/n) \qquad (m = 0, 1, 2, \cdots, r).$$

This follows at once if in (14.03) we put $j = m + 2$, set

$$T_n \equiv T_n^{(m)} \equiv \sum_{k=1}^{n} k(n-k)^m|\Delta^{m+1}p_k| = O(P_n), \qquad T_0 \equiv 0,$$

and proceed as in Lemma 8.1, using (16.02).

In virtue of (3.10) and (3.11) we write

$$(16.04)\qquad \mathfrak{N}_n(t) \equiv \tilde{N}_n(t) + iN_n(t) \equiv \frac{-e^{-i(n+\frac{1}{2})t}}{2\pi P_n \sin\frac{1}{2}t}\sum_{k=0}^{n} p_k e^{ikt}$$

and denote by $\mathfrak{N}_n^{(m)}(t)$ the result of differentiating $\mathfrak{N}_n(t)$ m times with respect to t.

Lemma 16.1. *For $0 \leqq t \leqq n^{-1}$, we have*

$$t^r N_n^{(r)}(t) = O(n).$$

First, for the values of t indicated and $k \leqq n$,

$$\begin{aligned}\left|t^r\frac{d^r}{dt^r}\frac{\sin(k+\frac{1}{2})t}{\sin\frac{1}{2}t}\right| &\leqq t^r\left\{\left|\frac{d^r}{dt^r}\cot\tfrac{1}{2}t\sin kt\right| + \left|\frac{d^r}{dt^r}\cos kt\right|\right\}\\ &\leqq t^r\left|\sum_{\beta=0}^{r-1}\binom{r}{\beta}\frac{d^\beta}{dt^\beta}\cot\tfrac{1}{2}t\,\frac{d^{r-\beta}}{dt^{r-\beta}}\sin kt + \sin kt\frac{d^r}{dt^r}\cot\tfrac{1}{2}t\right| + (kt)^r\\ &= O\left(t^r\sum_{\beta=0}^{r-1}\frac{1}{t^{\beta+1}}k^{r-\beta} + kt^{r+1}\cdot\frac{1}{t^{r+1}}\right) = O(k),\end{aligned}$$

whence by (3.10) and (2.02),

$$|t^r N_n^{(r)}(t)| = O\left(\frac{1}{|P_n|}\sum_{k=0}^{n}k|p_{n-k}|\right) = O(n).$$

To simplify the succeeding estimates we set

$$(16.05)\qquad \frac{1}{b_j(t)} \equiv 2\pi P_n(1-e^{it})^j\sin\tfrac{1}{2}t \qquad (|t| > 0).$$

Lemma 16.2. *We have*

$$\frac{d^h}{dt^h}b_j(t) = O\left(\frac{1}{|P_n|\,t^{j+h+1}}\right).$$

Lemma 16.3. *If $\{p_\nu\}$ satisfies* (16.01) *and* (16.03), *then*

$$t^m\mathfrak{N}_n^{(m)}(t) = o(1) \qquad (m = 0, 1, 2, \cdots, r)$$

uniformly in t for $0 < \delta \leqq |t| \leqq \pi$, where δ is fixed.

We first transform the sum in (16.04). Assuming $n > r + 1$, and applying Abel's partial summation formula as in §9 $s = m + 1$ times, we have

$$\sum_{k=0}^{n} p_k e^{ikt} = \sum_{k=0}^{n} \frac{(\Delta^s p_k)e^{ikt}}{(1 - e^{it})^s} - e^{i(n+1)t} \sum_{j=1}^{s} \frac{\Delta^{j-1} p_n}{(1 - e^{it})^j},$$

the differences being given by (14.01). Substituting into (16.04) and using the notation of (16.05) give

$$\mathfrak{N}_n(t) = -b_s(t) \sum_{k=0}^{n} (\Delta^s p_k) e^{-i(n-k+\frac{1}{2})t} + e^{\frac{1}{2}it} \sum_{j=1}^{s} b_j(t)\, \Delta^{j-1} p_n,$$

whence applying Lemma 16.2,

$$\begin{aligned}
\mathfrak{N}_n^{(m)}(t) &= -\sum_{\beta=0}^{m} \binom{m}{\beta} \Bigg[\frac{d^{m-\beta}}{dt^{m-\beta}} b_s(t) \sum_{k=0}^{n} (\Delta^s p_k) \frac{d^\beta}{dt^\beta} e^{-i(n-k+\frac{1}{2})t} \\
&\qquad - \frac{d^{m-\beta}}{dt^{m-\beta}} e^{\frac{1}{2}it} \sum_{j=1}^{s} (\Delta^{j-1} p_n) \frac{d^\beta}{dt^\beta} b_j(t)\Bigg] \\
&= O\left\{\sum_{\beta=0}^{m} \frac{1}{|P_n|} \left[\sum_{k=0}^{n} \frac{|\Delta^s p_k|\,(n-k)^\beta}{t^{s+m-\beta+1}} + \sum_{j=1}^{s} \frac{|\Delta^{j-1} p_n|}{t^{j+\beta+1}}\right]\right\} \\
&= \begin{cases} O\left(\dfrac{1}{|P_n|} \displaystyle\sum_{k=0}^{n} (n-k)^m |\Delta^s p_k| + \dfrac{1}{|P_n|} \displaystyle\sum_{j=1}^{s} |\Delta^{j-1} p_n|\right) & (t \geqq 1), \\[2ex] O\left(\displaystyle\sum_{k=0}^{n} \dfrac{(n-k)^m |\Delta^s p_k|}{|P_n|\, t^{s+m+1}} + \displaystyle\sum_{j=1}^{s} \dfrac{|\Delta^{j-1} p_n|}{|P_n|\, t^{j+m+1}}\right) & (t \leqq 1). \end{cases}
\end{aligned}$$

Now for $1 \leqq t \leqq \pi$, we have the desired result at once upon using the hypotheses, and for $\delta \leqq t \leqq 1$,

$$t^m \mathfrak{N}_n^{(m)}(t) = O\left(\frac{1}{|P_n|} \sum_{k=0}^{n} (n-k)^m |\Delta^s p_k| + \frac{1}{|P_n|} \sum_{j=1}^{s} |\Delta^{j-1} p_n|\right) = o(1).$$

This completes the proof of the lemma.

LEMMA 16.4. *If* $\{p_\nu\}$ *satisfies* (16.01) *and* (16.03), *then*

$$\mathfrak{N}_n^{(m)}(t) = o(1) \qquad (m = 0, 1, 2, \cdots, r)$$

uniformly in t *for* $0 < \delta \leqq |t| \leqq \pi$.

This follows at once from the previous lemma.

It remains to give estimates for the kernel in the interval (n^{-1}, δ). We shall use the notation of (9.05) and (9.06), and for $\beta = 0, 1, 2, \cdots, r$ put

$$W_m^\beta \equiv \sum_{k=0}^{m} (n-k)^\beta |\Delta^{r+1} p_k|, \qquad W^\beta(u) \equiv W_{[u]}^\beta. \tag{16.06}$$

Consider

$$\sum_{k=0}^{n} p_k e^{ikt} = \Sigma_1 + \Sigma_2 \qquad (t > 0), \tag{16.07}$$

where in Σ_1 k ranges over the integers $\leqq \tau = [t^{-1}]$, and in Σ_2 over the integers $> \tau$ but $\leqq n$.

The left member of (16.07) has continuous derivatives of all orders. Each term on the right has derivatives of all orders everywhere except at a denumerable set of points, viz., those for which $\tau = [t^{-1}]$, $(n^{-1} \leqq t \leqq \delta)$. But upon multiplying the r-th derivatives of the left and right members by a bounded function and integrating, we find the results are equal. Hence τ may be considered a constant with respect to differentiation.

We proceed to transform Σ_2, applying Abel's partial summation formula as in estimating the Σ_2 of (9.08). Assuming $n > r + 1$ and repeating $s = r + 1$ times, we get

$$\Sigma_2 = \sum_{j=1}^{s} \frac{(\Delta^{j-1} p_\tau)e^{i(\tau+1)t} - (\Delta^{j-1} p_n)e^{i(n+1)t}}{(1 - e^{it})^j} + \sum_{k=\tau+1}^{n} \frac{(\Delta^s p_k)e^{ikt}}{(1 - e^{it})^s},$$

where the differences are given by (14.01). Substituting this into (16.07) and the result into (16.04), we have

$$\begin{aligned} \mathfrak{N}_n(t) = \frac{-1}{2\pi P_n \sin \frac{1}{2}t} \Big[& \sum_{k=0}^{\tau} p_k e^{-i(n-k+\frac{1}{2})t} - \sum_{j=1}^{s} (\Delta^{j-1} p_\tau) b_j(t) e^{-i(n-\tau+\frac{1}{2})t} \\ & + \sum_{j=1}^{s} (\Delta^{j-1} p_n) b_j(t) e^{\frac{1}{2}it} - b_s(t) \sum_{k=\tau+1}^{n} (\Delta^s p_k) e^{-i(n-k+\frac{1}{2})t} \Big] \end{aligned}$$

$$\equiv L_{n1}(t) + L_{n2}(t) + L_{n3}(t) + L_{n4}(t). \tag{16.08}$$

We compute now the r-th derivative of each $L_{nm}(t)$ using the notation of (9.05), (9.06), and (16.06), together with Lemma 16.2, assuming that t^{-1} is not an integer.

For $m = 1$,

$$\begin{aligned} \frac{d^r}{dt^r} L_{n1}(t) &= \frac{-1}{2\pi P_n} \sum_{\beta=0}^{r} \binom{r}{\beta} \sum_{k=0}^{\tau} p_k \frac{d^\beta}{dt^\beta} e^{-i(n-k+\frac{1}{2})t} \frac{d^{r-\beta}}{dt^{r-\beta}} \frac{1}{\sin \frac{1}{2}t} \\ &= O\left(\frac{1}{|P_n|} \sum_{\beta=0}^{r} \sum_{k=0}^{\tau} (n-k)^\beta |p_k| t^{\beta-r-1}\right) = O\left(\frac{1}{|P_n|} \sum_{\beta=0}^{r} n^\beta t^{\beta-r-1} \sum_{k=0}^{\tau} r_k\right) \\ &= O\left(\frac{1}{|P_n|} R\left(\frac{1}{t}\right) \sum_{\beta=0}^{r} n^\beta t^{\beta-r-1}\right). \end{aligned}$$

For $m = 2$, putting $s = r + 1$, we get

$$\begin{aligned} \frac{d^r}{dt^r} L_{n2}(t) &= -\sum_{j=1}^{r+1} \left[\sum_{\beta=0}^{r} \binom{r}{\beta} (\Delta^{j-1} p_\tau) \frac{d^\beta}{dt^\beta} e^{-i(n-\tau+\frac{1}{2})t} \frac{d^{r-\beta}}{dt^{r-\beta}} b_j(t) \right] \\ &= O\left(\sum_{j=1}^{r+1} \sum_{\beta=0}^{r} \frac{|\Delta^{j-1} p_\tau| (n-\tau)^\beta}{|P_n| t^{j+r-\beta+1}} \right). \end{aligned}$$

For $m = 3$, we have at once

$$\frac{d^r}{dt^r} L_{n3}(t) = O\left(\sum_{j=1}^{r+1} \frac{|\Delta^{j-1} p_n|}{|P_n| t^{j+r+1}} \right).$$

Finally, for $m = 4$,

$$\frac{d^r}{dt^r} L_{n4}(t) = -\sum_{\beta=0}^{r} \binom{r}{\beta} \sum_{k=\tau+1}^{n} (\Delta^s p_k) \frac{d^\beta}{dt^\beta} e^{-i(n-k+\frac{1}{2})t} \frac{d^{r-\beta}}{dt^{r-\beta}} b_s(t)$$

$$= O\left(\sum_{\beta=0}^{r} \sum_{j=\tau+1}^{n} \frac{(n-k)^\beta \,|\, \Delta^s p_k \,|}{|\, P_n \,|\, t^{s+r-\beta+1}}\right) = O\left(\sum_{\beta=0}^{r} \frac{W^\beta(n) - W^\beta(t^{-1})}{|\, P_n \,|\, t^{s+r-\beta+1}}\right).$$

If we refer now to (16.04) and (16.08), these estimates imply the existence of a generic constant A such that

$$|\, N_n^{(r)}(t) \,|,\ |\, \tilde{N}_n^{(r)}(t) \,| \leqq |\, \mathfrak{N}_n^{(r)}(t) \,| \leqq A \sum_{m=1}^{4} M_{nm}(t),$$

where

$$M_{n1}(t) = \frac{1}{t^r R(n)} \sum_{\beta=0}^{r} \frac{n^\beta}{t^{-\beta+1}} R(t^{-1}), \qquad M_{n2}(t) = \frac{1}{t^r R(n)} \sum_{j=1}^{r+1} \sum_{\beta=0}^{r} \frac{|\, \Delta^{j-1} p_\tau \,|\, (n-\tau)^\beta}{t^{j-\beta+1}},$$

$$M_{n3}(t) = \frac{1}{t^r R(n)} \sum_{j=1}^{r+1} \frac{|\, \Delta^{j-1} p_n \,|}{t^{j+1}}, \qquad M_{n4}(t) = \frac{1}{t^r R(n)} \sum_{\beta=0}^{r} \frac{W^\beta(n) - W^\beta(t^{-1})}{t^{r-\beta+2}}.$$

17. **Proof of Theorem** III. The truth of Theorem III now follows from the following lemmas.

LEMMA 17.1. *As $n \to \infty$, $U_i(r, n) \to 0$.*

This follows at once from Lemma 16.4 and formula (13.06).

LEMMA 17.2. *If at $u = x$, $f^{(r)}(u)$ exists, then as $n \to \infty$,*

$$\int_0^{n^{-1}} t^r \omega(x, t)\, N_n^{(r)}(t) dt = o(1).$$

By Definition 13.1 and Lemma 16.1,

$$\int_0^{n^{-1}} t^r \omega(x, t)\, N_n^{(r)}(t) dt = O\left(n \int_0^{n^{-1}} |\, \omega(x, t) \,|\, dt\right) = o\left(n \int_0^{n^{-1}} dt\right) = o(1).$$

LEMMA 17.3. *If $\{p_\nu\}$ satisfies (16.01) and (16.03), then for fixed $\delta > 0$,*

$$\int_\delta^\pi t^r \omega(x, t)\, N_n^{(r)}(t) dt = o(1).$$

This follows from Definition 13.1 and Lemma 16.3 with $m = r$.

Thus (15.01) is reduced to

$$N_n^{(r)}[f(x), p_\nu] - f^{(r)}(x) = 2\,(-1)^r \int_{n^{-1}}^{\delta} \frac{t^r}{r!}\, \omega(x, t)\, N_n^{(r)}(t) dt + o(1),$$

and to complete the proof of Theorem III it suffices to show that

$$\int_{n^{-1}}^{\delta} t^r \,|\, \mathfrak{N}_n^{(r)}(t) \,|\, dt < C, \qquad 1/n < \delta,$$

as $\delta \to 0$ at the H_r regular points. This follows at once from

LEMMA 17.4. *If $\{p_\nu\}$ satisfies* (14.02), (14.03), *and* (14.04), *then*

$$\int_{n^{-1}}^{\delta} t^r M_{nm}(t)dt \leqq M \qquad (m = 1, 2, 3, 4). \tag{17.01}$$

Property (17.01) *holds in the cases* $m = 2, 3, 4$, *if we assume only that* (14.02) *and* (14.03) *are satisfied.*

For $m = 1$ we have

$$\int_{n^{-1}}^{\delta} t^r M_{n1}(t)dt = \sum_{\beta=0}^{r} \frac{n^\beta}{R(n)} \int_{n^{-1}}^{\delta} R(t^{-1})t^{\beta-1}dt = \sum_{\beta=0}^{r} \frac{n^\beta}{R(n)} \int_{\delta^{-1}}^{n} \frac{R(u)}{u^{\beta+1}}\, du,$$

which is bounded since

$$\frac{n^\beta}{R_n} \sum_{k=1}^{n} \frac{R_k}{k^{\beta+1}}$$

is bounded for $\beta = 0, 1, 2, \cdots, r$ by (16.02).

For $m = 2$,

$$\begin{aligned}\int_{n^{-1}}^{\delta} t^r M_{n2}(t)dt &= \sum_{j=1}^{r+1} \sum_{\beta=0}^{r} \frac{1}{R(n)} \int_{n^{-1}}^{\delta} \frac{(n-\tau)^\beta \mid \Delta^{j-1}p_\tau \mid}{t^{j-\beta+1}}\, dt \\ &= \sum_{j=1}^{r+1} \sum_{\beta=0}^{r} \frac{1}{R(n)} \int_{\delta^{-1}}^{n} u^{j-\beta-1}\, (n - [u])^\beta \mid \Delta^{j-1}p_{[u]} \mid du \\ &= O\left(\sum_{j=1}^{r+1} \sum_{\beta=0}^{r} \frac{1}{R_n} \sum_{k=1}^{n} k^{j-\beta-1}\, (n-k)^\beta \mid \Delta^{j-1}p_k \mid \right).\end{aligned}$$

This is bounded by (14.03).

For $m = 3$,

$$\int_{n^{-1}}^{\delta} t^r M_{n3}(t)dt = \sum_{j=1}^{r+1} \frac{\mid \Delta^{j-1}p_n \mid}{R_n} \int_{n^{-1}}^{\delta} \frac{dt}{t^{j+1}} = O\left(\sum_{j=1}^{r+1} \frac{n^j \mid \Delta^{j-1}p_n \mid}{R_n} \right).$$

This is bounded by (14.02).

Finally, for $m = 4$ we have

$$\begin{aligned}\int_{n^{-1}}^{\delta} t^r M_{n4}(t)dt &= \sum_{\beta=0}^{r} \frac{1}{R(n)} \int_{n^{-1}}^{\delta} \frac{W^\beta(n) - W^\beta(t^{-1})}{t^{r-\beta+2}}\, dt \\ &= \sum_{\beta=0}^{r} \frac{1}{R(n)} \int_{\delta^{-1}}^{n} u^{r-\beta}[W^\beta(n) - W^\beta(u)]\, du \\ &< \sum_{\beta=0}^{r} \frac{1}{R(n)} \int_{0}^{n} u^{r-\beta}[W^\beta(n) - W^\beta(u)]\, du \\ &= \sum_{\beta=0}^{r} \frac{1}{R(n)\,(r-\beta+1)} \int_{0}^{n} u^{r-\beta+1}\, dW^\beta(u) \\ &= O\left(\sum_{\beta=0}^{r} \frac{1}{R_n} \sum_{k=1}^{n} k^{r-\beta+1}(n-k)^\beta \mid \Delta^{r+1}p_k \mid \right).\end{aligned}$$

This is bounded by (14.03).

The proof of Theorem III is thus completed.

18. **Lemmas for Theorem** IV. We give here special lemmas for Theorem IV. Proofs are omitted, since they depend on elementary computations.

LEMMA 18.1. *For* $0 \leq t \leq n^{-1}$ *and* $k \leq n$, $H_k^{(r)}(t) \Big/ \dfrac{d^r}{dt^r} \cot \tfrac{1}{2}t = O(1)$.

LEMMA 18.2. *For* $0 \leq t \leq n^{-1}$ *and* $k \leq n$,

$$\frac{d}{dt}\left[H_k^{(r)}(t) \Big/ \frac{d^r}{dt^r} \cot \tfrac{1}{2}t\right] = O(n).$$

LEMMA 18.3. *At* $t = n^{-1}$ *and for* $0 \leq k \leq n$,

$$\frac{d^r}{dt^r} \frac{\cos (k + \frac{1}{2})t}{\sin \frac{1}{2}t} \Big/ \frac{d^r}{dt^r} \cot \tfrac{1}{2}t = O(1).$$

LEMMA 18.4. *If at* $u = x$, $\bar{f}^{(r)}(u)$ *exists, then as* $\alpha, \beta \to 0$,

$$\int_\alpha^\beta A_r(t) \frac{d^r}{dt^r} \cot \tfrac{1}{2}t \, dt \to 0.$$

We note also, since the restrictions on $\{p_\nu\}$ in Theorem IV are the same as for Theorem III except that in the former we replace the r of the latter by $r + 1$, that all the lemmas and results of the previous two sections may be used here with the index r increased by unity.

19. **Proof of Theorem** IV. The truth of the theorem now follows from the following lemmas.

LEMMA 19.1. *As* $n \to \infty$, $D_r(n) \to 0$.

This follows at once from the expressions (3.11), (13.05) and (16.04), together with Lemma 16.4.

LEMMA 19.2. *As* $n \to \infty$, $J_1 \to 0$.

Integrating by parts and using Lemmas 18.1, 18.2 and 18.4, together with the notation of (3.11) and (13.02), we have from (15.02) that

$$\begin{aligned} J_1 &= \frac{(-1)^{r+1}}{2\pi P_n} \sum_{k=0}^{n} p_{n-k} \int_0^{n^{-1}} A_r(t)\, H_k^{(r)}(t)\, dt \\ &= \frac{(-1)^{r+1}}{2\pi P_n} \sum_{k=0}^{n} p_{n-k} \left[- A_r^*\left(\frac{1}{n}\right) \frac{H_k^{(r)}(t)}{\dfrac{d^r}{dt^r} \cot \frac{1}{2}t} \Bigg|_0^{n^{-1}} + \int_0^{n^{-1}} A_r^*\left(\frac{1}{n}\right) \frac{d}{dt} \frac{H_k^{(r)}(t)}{\dfrac{d^r}{dt^r} \cot \frac{1}{2}t}\, dt \right] \\ &= o\left(\frac{1}{|P_n|} \sum_{k=0}^{n} |p_{n-k}|\right) + o\left(\frac{1}{|P_n|} \sum_{k=0}^{n} |p_{n-k}| \int_0^{n^{-1}} k\, dt\right) = o(1), \end{aligned}$$

in virtue of (2.02).

LEMMA 19.3. *As* $n \to \infty$, $\delta > 0$ *being fixed*, $J_3 \to 0$.

This follows at once from (15.02), since by (16.04) and Lemma 16.4 the kernel of J_3 tends to zero uniformly in t, $0 < \delta \leq t \leq \pi$, as $n \to \infty$.

LEMMA 19.4. *As* $\delta \to 0$, $J_2 \to 0$.

Integrating by parts and using the notation of (13.02), we have from (15.02) that

$$J_2 = -A_r^*(\delta) \frac{\bar{N}_n^{(r)}(t)}{\dfrac{d^r}{dt^r} \cot \frac{1}{2}t} \Bigg|_{n^{-1}}^{\delta} + \int_{n^{-1}}^{\delta} A_r^*(\delta) \frac{d}{dt}\left[\bar{N}_n^{(r)}(t) \Big/ \frac{d^r}{dt^r} \cot \tfrac{1}{2}t\right] dt$$

$$= o(1) + o\left(\int_{n^{-1}}^{\delta} (t^{r+1} \,|\, \mathfrak{N}_n^{(r+1)}(t) \,|\, + t^r \,|\, \mathfrak{N}_n^{(r)}(t) \,|) \, dt\right) = o(1),$$

first by using Lemmas 16.4, 18.3 and 18.4, together with (2.02), and then Lemma 17.4 under the hypotheses of Theorem IV.

This completes the proof of the theorem.

20. (C, α) **continuity.** We give here the theorem for K_α effectiveness of the method (N, p_ν) in the case $\alpha > 1$. Let r be an integer such that $1 \leqq r - 1 < \alpha \leqq r$.

THEOREM V. *A regular Nörlund method of summation* (N, p_ν) *is* K_α *effective* $(\alpha > 1)$, *provided that* $\{p_\nu\}$ *satisfies conditions* (14.02), (14.03) *and* (14.04).

Integrating (4.03) by parts r times, we have

$$\begin{aligned} N_n[f(x), p_\nu] - f(x) &= \int_0^\pi \varphi(t) N_n(t) \, dt \\ &= \sum_{m=0}^{r-1} (-1)^m \varphi_{m+1}(\pi) N_n^{(m)}(t) \Big|_{t=\pi} \\ &\qquad + (-1)^r \left[\int_0^\delta + \int_\delta^\pi\right] \varphi_r(t) N_n^{(r)}(t) \, dt \\ &\equiv \sum_{m=0}^{r-1} A_n^{(m)} + I_n' + I_n''. \end{aligned}$$

By Lemma 16.4 and (16.04), each $A_n^{(m)} \to 0$ as $n \to \infty$, and $I_n'' \to 0$ for fixed $\delta > 0$. Finally, since $\alpha \leqq r$, $\varphi_r(t) = o(t^r)$, so that as $n \to \infty$ and $\delta \to 0$,

$$|\, I_n' \,| = o\left(\int_0^\delta t^r \,|\, \mathfrak{N}_n^{(r)}(t) \,|\, dt\right) = o(1),$$

by an argument similar to that used in Lemma 17.2, and by Lemma 17.4.

IV. Strong summability

21. **Definitions and results.** Hardy and Littlewood [2] have shown that if for $p > 1$

$$\text{(21.01)} \qquad \int_0^t |\, \varphi(u) \,|^p \, du = O(t), \qquad \varphi_1(t) = o(t),$$

the Fourier series of $f(x)$ is strongly summable, i.e.,

$$\text{(21.02)} \qquad \sum_{m=0}^{n} |\, s_m(x) - f(x) \,|^q = o(n)$$

for every positive q; and if

$$\int_0^t |\psi(u)|^p\, du = O(t), \tag{21.03}$$

then at points where $\bar{f}(x)$ has a definite value, the conjugate series is also strongly summable. Previously, O. G. Sutton [14] had proved that (21.02) holds if we replace $s_n(x)$ by its n-th Fejér mean. R. E. A. C. Paley [9][4] showed that if

$$\int_0^t |\varphi(u)|\,\{\log^+|\varphi(u)|\}^{1+\delta} = O(t), \qquad \varphi_k(t) = o(t^k), \tag{21.04}$$

where $\delta > 0$ and k is any positive number, then (21.02) holds.

From the results of C. E. Winn [16], it follows that for $q = 1$ (21.02) holds if we replace $s_n(x)$ by its Cesàro transform of any positive order, assuming (21.01) or (21.04); and that the conjugate Fourier series is strongly summable for $q = 1$ if we replace $\bar{s}_n(x)$ by its Cesàro transform of any order assuming (21.03). Our purpose is to give theorems on strong summability for the method (N, p_ν). We set $r = q/(1 + \eta)$ for fixed positive $\eta < 1$, and denote by $\{N_n[s_k]\}$, $\{N_n[\bar{s}_k]\}$ the Nörlund transforms of $\{s_k(x)\}$, $\{\bar{s}_k(x)\}$, respectively. Our theorems are the following:

THEOREM VI. *If $\varphi(t)$ satisfies* (21.04), *then for every positive q*

$$\sum_{m=0}^{n} |N_m[s_k] - f(x)|^q = o(n),$$

provided the regular generating sequence $\{p_\nu\}$ satisfies the condition[5]

$$\sum_{k=0}^{m} \left|\frac{p_k}{P_m}\right|^{r'} = O(m^{1-r'}). \tag{21.05}$$

THEOREM VII. *The conjugate Fourier series of $f(x)$ is strongly summable at all points where $\bar{f}(x)$ has a definite finite value and*

$$\int_0^t |\psi(u)|\,\{\log^+|\psi(u)|\}^{1+\delta}\, du = O(t) \qquad (\delta > 0). \tag{21.06}$$

THEOREM VIII. *If $\psi(t)$ satisfies* (21.06), *then for every positive q*

$$\sum_{m=0}^{n} |N_m[\bar{s}_k] - \bar{f}(x)|^q = o(n)$$

at all points where $\bar{f}(x)$ has a definite finite value, provided the regular generating sequence $\{p_\nu\}$ satisfies (21.05).

22. **Proof of Theorem** VI. An application of Hölders inequality shows that we may assume $q > 2$ and therefore $r > 1$. We assume with Paley that the standard simplifications have been made so that $f(t) \sim \Sigma a_n \cos nt$ is an even function with zero mean value, and $x = 0$, $f(0) = 0$, so that $s_m(0) \equiv s_m = a_1 + a_2 + \cdots + a_m$.

[4] All future page references are to this paper.

[5] For any positive number s we define s' by the relation $1/s + 1/s' = 1$.

Let us put $\rho = \exp(-\delta/n)$, where $0 < \delta < 1$. Then[6]

$$\bar{s}_k \equiv \sum_{\nu=1}^{k} a_\nu \rho^\nu = s_k \rho^k + (1-\rho) \sum_{\nu=1}^{k-1} s_\nu \rho^\nu = s_k \rho^k + O\left(\frac{\delta m}{n}\right) \qquad (m \leqq n),$$

where the "O" is independent of k, m, and n. In virtue of the regularity condition (2.02)

$$|N_m[\bar{s}_k] - N_m[s_k \rho^k]| \leqq A \frac{\delta m}{n} \leqq A\delta \qquad (m \leqq n),$$

where A is a generic constant, so that if $\epsilon < 1$ is a preassigned positive number, we choose δ so small that

$$\sum_{m=1}^{n} |N_m[\bar{s}_k] - N_m[s_k \rho^k]|^q \leqq \epsilon n. \tag{22.01}$$

Now

$$\bar{s}_k = \frac{1}{\pi}\int_0^\pi \frac{u(\rho,\theta) \sin(k+\frac{1}{2})\theta}{2\sin\frac{1}{2}\theta}\,d\theta = \frac{1}{\pi}\int_0^\pi u(\rho,\theta)\cot\tfrac{1}{2}\theta \sin k\theta d\theta + \frac{1}{\pi}\int_0^\pi u(\rho,\theta)\cos k\theta d\theta, \tag{22.02}$$

where $u(r, \theta)$ is the Poisson integral of $f(t)$.

By the Riemann-Lebesgue theorem the last term on the right in (22.02) is $o(1)$ as $k \to \infty$. Hence in virtue of the regularity of the transformation, its Nörlund limit is zero. The first term we write as

$$\int_0^{G(1-\rho)} + \int_{G(1-\rho)}^{\pi} \equiv s_{k1} + s_{k2},$$

where G remains to be chosen. Then by Minkowski's inequality,

$$\left(\sum_{m=1}^{n} |N_m[\bar{s}_k]|^q\right)^{\frac{1}{q}} \leqq \left(\sum_{m=1}^{n} |N_m[s_{k1}]|^q\right)^{\frac{1}{q}} + \left(\sum_{m=1}^{n} |N_m[s_{k2}]|^q\right)^{\frac{1}{q}} + o\left(n^{\frac{1}{q}}\right). \tag{22.03}$$

Now[7] $|s_{k1}| \leqq \epsilon$ $(k \leqq n)$ for any fixed G, so that by (2.02),

$$\sum_{m=1}^{n} |N_m[s_{k1}]|^q \leqq (C\epsilon)^q n. \tag{22.04}$$

By Hölder's inequality and (21.05),

$$\left|\sum_{k=1}^{m} \frac{p_{m-k}}{P_m} s_{k2}\right|^r \leqq \left(\sum_{k=1}^{m} \left|\frac{p_{m-k}}{P_m}\right|^{r'}\right)^{\frac{r}{r'}} \left(\sum_{k=1}^{m} |s_{k2}|^r\right) \leqq A\,\frac{1}{m} \sum_{k=1}^{m} |s_{k2}|^r.$$

Hence by a lemma due to Hardy [1],

$$\sum_{m=1}^{n} |N_m[s_{k2}]|^{r(1+\eta)} \leqq A\left(\frac{1}{m}\sum_{k=1}^{m} |s_{k2}|^r\right)^{1+\eta} \leqq A \sum_{k=1}^{m} |s_{k2}|^q \leqq A\epsilon^{q-1} n$$

[6] Loc. cit., p. 433, (4.2).

[7] Loc. cit., p. 437.

by proper choice[8] of G. Substituting into (22.03) and using (22.04), we have

$$\left(\sum_{m=1}^{n} |N_m[\bar{s}_k]|\right)^{\frac{1}{q}} \leqq n^{\frac{1}{q}}\left\{C\epsilon + A\epsilon^{\frac{q-1}{q}}\right\} + o\left(n^{\frac{1}{q}}\right) = o\left(n^{\frac{1}{q}}\right).$$

Hence, again using Minkowski's inequality, we have finally from (22.01) together with the fact that $\rho^k \geqq e^{-1}$ for $k \leqq n$

$$\sum_{m=1}^{n} |N_m[s_k]|^q = o(n).$$

This completes the proof of the theorem.

23. **Proof of Theorems** VII **and** VIII. We shall first prove Theorem VII. As before, we assume the standard simplifications have been made, i.e., $f(x)$ is an odd function, $x = 0$ and $\bar{f}(0) = 0$. The conjugate series is then $-\Sigma b_n$ and $\bar{s}_m(0) \equiv \sigma_m = -(b_1 + b_2 + \cdots + b_m)$.

LEMMA 23.1. *Let τ_m denote the m-th Cesàro sum of the conjugate Fourier series of $f(x)$ at $x = 0$. Then τ_m is uniformly bounded.*

From (21.06),

$$\int_0^t |\psi(u)|\, du = O(t).$$

The rest of the proof is similar to that of Fejér's theorem.

As before,

$$\bar{\sigma}_m \equiv -\sum_{\nu=1}^{m} b_\nu \rho^\nu = \sigma_m \rho^m + O\left(\frac{\delta m}{n}\right) \qquad (m \leqq n),$$

so that by proper choice of δ,

$$\sum_{m=1}^{n} |\bar{\sigma}_m - \sigma_m \rho^m| \leqq \epsilon n. \tag{23.01}$$

Now

$$\begin{aligned}\bar{\sigma}_m &= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\cos(m+\frac{1}{2})\theta - \cos\frac{1}{2}\theta}{2\sin\frac{1}{2}\theta} u(\rho, \theta)\, d\theta \\ &= \frac{1}{\pi}\int_0^{\pi} u(\rho,\theta)\cos m\theta \cot\tfrac{1}{2}\theta\, d\theta - \frac{1}{\pi}\int_0^{\pi} u(\rho,\theta)\sin m\theta\, d\theta \\ &\qquad - \frac{1}{\pi}\int_0^{\pi} u(\rho,\theta)\cot\tfrac{1}{2}\theta\, d\theta.\end{aligned}$$

The second term on the right is $o(1)$ by the Riemann-Lebesgue theorem. The last term, as $\rho \to 1$, is $\bar{f}(0)$, which is zero by hypothesis. Hence we may write

$$\bar{\sigma}_m = \frac{1}{\pi}\left(\int_0^{G(1-\rho)} + \int_{G(1-\rho)}^{\pi}\right) u(\rho,\theta)\cos m\theta\cot\tfrac{1}{2}\theta\, d\theta + o(1)$$

[8] Loc. cit., p. 436, (7.2).

$$\equiv \sigma_{m1} + \sigma_{m2} + o(1), \tag{23.02}$$

where G remains to be chosen.

We consider first the term σ_{m2} and observe that it is essentially the m-th Fourier cosine coefficient of the even function $h(t)$ which is zero in the interval $[0, G(1 - \rho)]$ and in $[G(1 - \rho), \pi]$ is equal to $u(\rho, t)\cot \frac{1}{2}t$. Hence applying the Young-Hausdorff theorem we have as in the paper of Paley[9]

$$\sum_{m=1}^{n} | \sigma_{m2} |^q \leqq \epsilon^{q-1} n. \tag{23.03}$$

For σ_{m1}, we observe that in the interval $[0, G(1 - \rho)]$ the function $\cos mt$, since $m \leqq n$, has no more than AG turning points. Hence this interval can be divided into less than AG subintervals in each of which $\cos mt$ is monotonic. Applying the second law of the mean to the integral over each such subinterval, $| \sigma_{m1} |$ is not greater than the sum of AG terms of the form

$$\left| \int_{\xi}^{\xi'} u(\rho, \theta) \cot \tfrac{1}{2}\theta \, d\theta \right|,$$

where $0 \leqq \xi < \xi' \leqq G(1 - \rho)$. Hence for fixed G and $n \to \infty$, we have

$$\sum_{m=1}^{n} | \sigma_{m1} |^q \leqq \epsilon^q n. \tag{23.04}$$

Returning now to (23.02), we have upon applying Minkowski's inequality, and using (23.03) and (23.04), that

$$\left(\sum_{m=1}^{n} | \bar{\sigma}_m |^q\right)^{\frac{1}{q}} \leqq n^{\frac{1}{q}}\left(\epsilon + \epsilon^{\frac{q-1}{q}}\right) + o\left(n^{\frac{1}{q}}\right) = o\left(n^{\frac{1}{q}}\right).$$

Hence by (23.01) and again using Minkowski's inequality,

$$\left(\sum_{m=1}^{n} | \sigma_m \rho^m |^q\right)^{\frac{1}{q}} \leqq \left(\sum_{m=1}^{n} | \bar{\sigma}_m |^q\right)^{\frac{1}{q}} + \left(\sum_{m=1}^{n} | \bar{\sigma}_m - \sigma_m \rho^m |^q\right)^{\frac{1}{q}} = o\left(n^{\frac{1}{q}}\right).$$

Finally, then, since for $m \leqq n$ we have $\rho^m \geqq e^{-1}$, it follows that

$$\sum_{m=1}^{n} | \sigma_m |^q = o(n).$$

This completes the proof of Theorem VII.

The proof of Theorem VIII is similar to that of Theorem VI.

BIBLIOGRAPHY

1. G. H. Hardy, *Note on a theorem of Hilbert*, Mathematische Zeitschrift, vol. 6 (1920), pp. 314–317.
2. G. H. Hardy and J. E. Littlewood, *Notes on the theory of series* (IV): *On the strong summability of Fourier series*, Proceedings of the London Mathematical Society, (2), vol. 26 (1927), pp. 273–286.

[9] Loc. cit., p. 436.

3. E. Hille and J. D. Tamarkin, *On the summability of Fourier series.* I, Transactions of the American Mathematical Society, vol. 34 (1932), pp. 757–783.

4. M. Jacob, *Über die Cesàro'sche Summierbarkeit des Fourier'schen Integrales*, Bulletin International, Classe des Sciences Mathématiques et Naturelles, Cracow Academy of Science (A), 1926, pp. 41–74.

5. Ch. J. de la Vallée Poussin, *Sur l'approximation des fonctions d'une variable réelle*, Académie Bruxelles, Bulletin, (1908), pp. 193–254.

6. A. F. Moursund, *On summation of derived series of the conjugate Fourier series*, Annals of Mathematics, vol. 36 (1935), pp. 182–193.

7. A. F. Moursund, *On the r-th derived conjugate function*, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 131–136.

8. N. E. Nörlund, *Sur une application des fonctions permutables*, Lunds Universitets Årsskrift, (N.F.), avd. 2, 16, No. 3 (1920); 10 pp.

9. R. E. A. C. Paley, *On the strong summability of Fourier series*, Proceedings of the Cambridge Philosophical Society, vol. 26 (1930), pp. 429–437.

10. A. Plessner, *Zur Theorie der konjugierten trigonometrischen Reihen*, Mitteilungen des Mathematischen Seminars der Universität Giessen, Heft 10 (1923), 36 pp.

11. A. Plessner, *Trigonometrische Reihen*, in Pascal's *Repertorium der höheren Analysis*, vol. I_3, Leipzig und Berlin, 1929.

12. K. I. Sayers, *Cesàro summation of the differentiated series of Fourier-Lebesgue series and their allied series*, Proceedings of the London Mathematical Society, (2), vol. 31 (1930), pp. 29–39.

13. A. H. Smith, *On the summability of derived conjugate series of the Fourier-Lebesgue type*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 406–412.

14. O. G. Sutton, *On a theorem of Carleman*, Proceedings of the London Mathematical Society, (2), vol. 23 (1925), pp. xlviii–li.

15. T. Takahashi, *Note on the Cesàro summability of the conjugate series of the derived Fourier series*, Japanese Journal of Mathematics, vol. 10 (1933–34), pp. 127–132.

16. C. E. Winn, *On strong summability for any positive order*, Mathematische Zeitschrift, vol. 37 (1933), pp. 481–492.

17. G. F. Woronoi, *Extension of the notion of the limit of the sum of terms of an infinite series* (in Russian), Proceedings of the Eleventh Congress of Russian Naturalists and Physicians, St. Petersburg, 1902, pp. 60–61. Annotated English translation by J. D. Tamarkin, Annals of Mathematics, (2), vol. 33 (1932), pp. 422–428.

Brown University.

TWO SYSTEMS OF POLYNOMIALS FOR THE SOLUTION OF LAPLACE'S INTEGRAL EQUATION

By H. Bateman

1. In the radiation and conduction problems[1] in which the integral equation

$$f(x) = \int_0^\infty e^{-xt} g(t) dt$$

occurs, the variable x takes positive values, and so the function $g(t)$ is to be derived from the values of x for $x > 0$. In the inversion formulas given by Lord Kelvin[2]

$$f(x) = x^{-\frac{1}{2}} \int_0^\infty \frac{\cos}{\sin} (u/2x) F(u) du,$$

$$g(t) = (\pi^3 t)^{-\frac{1}{2}} \int_0^\infty du \int_0^\infty \frac{\text{ch}}{\text{sh}} (ut)^{\frac{1}{2}} \frac{\cos}{\sin} (ut)^{\frac{1}{2}} \frac{\cos}{\sin} (u/2x) x^{-1} f(x) dx,$$

the integration with respect to x does indeed run from 0 to ∞, but conditions to be satisfied by $f(x)$ or $F(u)$ sufficient to make one of these formulas valid have not yet been formulated in a useful form. A similar remark applies to the somewhat analogous formula of F. Sbrana.[3] A more complete inversion formula in which the integration runs from $x = 0$ to $x = \infty$ has been given recently by R. E. A. C. Paley and N. Wiener.[4] In Murphy's first method[5] of solving the integral equation, $xf(x)$ is expanded in a series of ascending powers of x^{-1} and $g(t)$ is expressed as the coefficient of x^{-1} in $f(x)e^{xt}$ which, by Cauchy's theory, may be expressed as a contour integral. This method was generalized by Lerch[6] for the case in which $x^{\nu}f(x)$ can be expanded in a series of powers of x^{-1} and the resulting expression can be transformed into a contour integral resembling that used in the well-known inversion formula of Laplace, Riemann and Mellin.

Murphy also gave a method in which $f(x)$ is expanded in a series of inverse

Received February 3, 1936; presented to the American Mathematical Society, November 30, 1935.

[1] H. Poincaré, Jour. de Phys., vol. 11 (1912), p. 34; L. Silberstein, Phil. Mag., vol. 15 (1932), p. 375; H. Bateman, Proc. Camb. Phil. Soc., vol. 15 (1910), pp. 423–427.

[2] Lord Kelvin, Camb. Math. Jour., vol. 3 (1842), p. 170; Math. and Phys. papers, vol. 1, p. 10. See also H. Bateman, Messenger of Math., vol. 57 (1928), p. 145.

[3] F. Sbrana, Rend. Lincei, (5), vol. 31 (1922), pp. 454–456.

[4] *Fourier Transforms in the Complex Domain*, Chapter 3.

[5] R. Murphy, Camb. Phil. Trans., vol. 4 (1833), p. 353.

[6] M. Lerch, Rozpravy, vol. 2 (1893), p. 9; Fortschritte der Math., vol. 25 (1893), p. 482.

factorials and $g(t)$ in a series of powers of $(1 - e^{-t})$; this method has been developed by Schlömilch[7] and is used now in discussions of factorial series.

In Murphy's third method,[8] $xf(x)$ is expanded in a series of powers of $1 - x^{-1}$ and $g(t)$ in a series of polynomials of Laguerre. This method has been used recently with some success by Picone[9] and Eckart [10] and the analysis developed further by Picone,[11] Tricomi[12] and Widder.[13] For the success of this method simple expressions for $f(x)$ and all its derivatives are needed for $x = \infty$. This is true also for a successful application of Widder's new method[14] which depends on the limiting form, as $n \to \infty$, of an expression involving the n-th derivative of $f(x)$ and furnishes a convenient way of approximating to $g(t)$ when the derivatives can be found with accuracy. When $f(x)$ is given by a graph, it is not advisable to use derivatives.[15]

2. In order to make use of integrals involving $f(x)$, it seems natural to adopt Murphy's idea of using an expansion in a series of orthogonal functions but to apply it to $f(x)$ instead of $g(t)$.[16] The simplest way of doing this is to use a set of orthogonal functions of type

$$\frac{1}{x+1} P_n\left(\frac{x-1}{x+1}\right),$$

where $P_n(z)$ is the Legendre polynomial. Writing

$$\frac{1}{x+1} P_n\left(\frac{x-1}{x+1}\right) = \int_0^\infty e^{-xt} U_n(t)dt, \tag{1}$$

the entire function $U_n(t)$ may be expressed in the form

$$U_n(t) = \frac{1}{2\pi}\int_0^{2\pi} \exp\left[-t(1 + e^{-i\theta})\right] P_n(1 + 2e^{i\theta})d\theta. \tag{2}$$

The power series for $U_n(t)$ is

$$U_n(t) = \sum_{m=0}^{\infty} (-1)^m \frac{t^m}{m!} F_n(-2m - 1), \tag{3}$$

[7] O. Schlömilch, Zeit. für Math. und Physik, vol. 4 (1859), p. 390.

[8] R. Murphy, Camb. Phil. Trans., vol. 5 (1835), p. 113. See especially p. 145.

[9] M. Picone, Rend. Sem. Mat., Univ. Roma, (1933).

[10] C. Eckart, Phys. Rev., (2), vol. 45 (1934), p. 851.

[11] M. Picone, Rend. Lincei, (6), vol. 21 (1935), p. 306.

[12] F. Tricomi, ibid., pp. 232, 420.

[13] D. V. Widder, this Journal, vol. 1 (1935), p. 126.

[14] D. V. Widder, Trans. Amer. Math. Soc., vol. 36 (1934), p. 107.

[15] Use can, however, be made of Widder's solution of the moment problem or some analogous formula in which use is made only of the values of $f(x)$ at a denumerable set of points.

[16] This plan has already been used by F. Sbrana, loc. cit. He used a Fourier series and a corresponding expansion of $g(t)$ in a series of Bessel functions of a type occurring in electromagnetic theory.

where $F_n(z)$ is a polynomial of degree n in z which has been studied elsewhere.[17] If $Q_n(z)$ denotes the Legendre function of the second kind, there is a Neumann series

$$\frac{1}{z-1}\exp\left[-t\frac{z+1}{z-1}\right] = \sum_{n=0}^{\infty}(2n+1)Q_n(z)U_n(t). \tag{4}$$

This may be proved by using the definite integral for $U_n(t)$ and the known properties of the expansion

$$\frac{1}{z-\mu} = \sum_{n=0}^{\infty}(2n+1)Q_n(z)P_n(\mu).$$

By equating coefficients of z^{-m-1} on the two sides of the equation (4), we obtain the expansion of a member of Laguerre's set of orthogonal functions

$$e^{-t}L_m(2t) = \sum_{n=0}^{m}(n+\tfrac{1}{2})U_n(t)\int_{-1}^{1}P_n(\mu)\mu^m d\mu. \tag{5}$$

This may be inverted, with the result

$$U_n(t) = e^{-t}\sum_{m=0}^{n}a_{n,m}L_m(2t) = e^{-t}Z_n(t), \tag{6}$$

where $a_{n,m}$ is the coefficient in the series

$$P_n(z) = \sum_{m=0}^{n}a_{n,m}z^m.$$

The functions $U_n(t)$ form an orthogonal set in a generalized sense for there is a relation

$$\int_0^\infty U_m(s)V_n(s)ds \equiv \int_0^\infty\int_0^\infty\frac{dsdt}{s+t}U_m(s)U_n(t) = \frac{\delta_{mn}}{2n+1}, \tag{7}$$

where the function $V_n(s)$ may be expressed in the following ways:

$$V_n(s) = \int_0^\infty\frac{dt}{s+t}U_n(t), \tag{8}$$

$$V_n(s) = \int_{-1}^{1}\frac{dz}{1-z}e^{-s\frac{1+z}{1-z}}P_n(z), \tag{9}$$

$$V_n(s) \sim \sum_{m=n}^{\infty}e^{-s}L_m(2s)\int_{-1}^{1}z^mP_n(z)dz. \tag{10}$$

It may also be defined as a coefficient in the Legendre series

$$\frac{1}{1-z}\exp\left[-s\frac{1+z}{1-z}\right] = \sum_{n=0}^{\infty}(n+\tfrac{1}{2})V_n(s)P_n(z) \qquad (s>0). \tag{11}$$

[17] H. Bateman, Tôhoku Math. Jour., vol. 37 (1933), p. 23; Annals of Math., vol. 35 (1934), p. 767.

To prove the relation (7), we write

$$(12) \qquad V_n(s) = \int_0^\infty \frac{dt}{s+t} e^{-t} \sum_{m=0}^{n} a_{n,m} L_m(2t) = \sum_{m=0}^{n} a_{n,m} W_m(s),$$

where

$$(13) \qquad W_m(s) = \int_0^\infty \frac{dt}{s+t} e^{-t} L_m(2t).$$

Consequently

$$\begin{aligned}\int_0^\infty U_m(s)\, W_p(s) ds &= \int_0^\infty U_m(s) ds \int_0^\infty \frac{dt}{s+t} e^{-t} L_p(2t) \\ &= \int_0^\infty dt\, e^{-t} L_p(2t)\, V_m(t) \\ &= \tfrac{1}{2} \int_{-1}^{1} z^p P_m(z) dz.\end{aligned}$$

To justify the change of order of integration in the repeated integral, we have merely to justify the change of order in the repeated integral

$$\int_0^\infty e^{-s} L_q(2s) ds \int_0^\infty \frac{dt}{s+t} e^{-t} L_p(2t),$$

for our repeated integral is the sum of a finite number of such terms. Each of these is, however, the sum of a finite number of terms of type

$$\int_0^\infty e^{-s} s^a ds \int_0^\infty \frac{dt}{s+t} e^{-t} t^b,$$

and in this repeated integral a change of order of integration is permissible because the integrand is positive.

To justify the value adopted for the integral

$$\int_0^\infty dt\, e^{-t} L_p(2t) V_m(t),$$

we have to verify the third expression for $V_n(s)$ and to show that this is indeed a Laguerre series. It should be noticed that $W_m(s)$ is of the form

$$A(s) W_0(s) + B(s),$$

where $A(s)$ and $B(s)$ are polynomials in s. Consequently $V_n(s)$ is also of this form. Now the expansion

$$(14) \qquad W_0(s) = \int_0^\infty \frac{dt}{s+t} e^{-t} = V_0(s) = \sum_{m=0}^{\infty} \frac{2}{2m+1} e^{-s} L_{2m}(2s) \qquad (s > 0)$$

is absolutely convergent because when m is large the m-th term is of order $m^{-\frac{5}{4}}$. It follows then, by Abel's theorem, that the series is equal to the limit when $t \to 1$ of the sum of the power series which represents

$$\int_{-t}^{t} \frac{dt_1}{1 - t_1} \exp\left[- s\frac{1 + t_1}{1 - t_1} \right],$$

and this is a Laguerre series in the variable s. The value on the left of (14) may thus be regarded as the sum of the series and the coefficients may be shown to be those derived by Laguerre's rule. It may now be shown by means of the recurrence relations

$$(n + 1)L_{n+1}(z) + nL_{n-1}(z) = (2n + 1 - z)L_n(z),$$

$$(n + 1)\; W_{n+1}\,(s) + n\; W_{n-1}\,(s) - (2n + 1 + 2s)\; W_n(s) + 2(-1)^n = 0$$

that

$$\int_0^\infty e^{-s}\; W_n(s)\; L_m\;(2s)\; ds = \tfrac{1}{2}\int_{-1}^{1} z^{m+n} dz.$$

Making use of (12) we have the desired relation which gives

$$\int_0^\infty U_m(s)V_n(s)ds = \sum_{p=0}^{n} a_{n,\,p} \int_0^\infty U_m(s)W_p(s)ds$$

$$= \tfrac{1}{2} \sum_{p=0}^{n} a_{n,\,p} \int_{-1}^{1} z^p P_m(z)dz = \tfrac{1}{2} \int_{-1}^{1} P_n(z)P_m(z)dz,$$

and so the formula (7) is established. It was derived originally from the orthogonal relation for the Legendre polynomials by using the expressions for these polynomials in terms of $U_m(s)$ and $U_n(t)$. When the integral equation possesses a suitable type of solution[18] $g(t)$, we may write

$$\int_0^\infty \frac{dx}{x+1} P_n\left(\frac{x-1}{x+1}\right) f(x) = \int_0^\infty f(x)dx \int_0^\infty e^{-xs}\; U_n(s)ds$$

$$= \int_0^\infty U_n(s)ds \int_0^\infty e^{-xs} f(x)dx$$

$$= \int_0^\infty U_n(s)ds \int_0^\infty dx \int_0^\infty e^{-x(s+t)}\, g(t)dt$$

$$= \int_0^\infty U_n(s)ds \int_0^\infty \frac{g(t)dt}{s + t}$$

$$= \int_0^\infty g(t)dt \int_0^\infty U_n(s)\, \frac{ds}{s+t} = \int_0^\infty g(t)V_n(t)dt.$$

[18] We make no attempt here to formulate conditions sufficient for the validity of these changes in the order of integration, though this can be done with the aid of the conditions of de La Vallée Poussin. See Bromwich, *Infinite Series*, p. 503.

Under these circumstances, if

$$f_m(x) = f(x) - \sum_{m=1}^{m} c_n \frac{1}{x+1} P_n\left(\frac{x-1}{x+1}\right),$$

$$g_m(t) = g(t) - \sum_{n=1}^{m} c_n U_n(t),$$

the coefficients c_n which make

$$\int_0^\infty [f_m(x)]^2 dx$$

as small as possible are the same as the coefficients c_n which make

$$\int_0^\infty g_m(t)dt \int_0^\infty \frac{g_m(s)ds}{s+t}$$

as small as possible and so we can form a definite idea of the way in which the present method gives an approximation to the solution of the integral equation.

There are, of course, cases in which the integral equation possesses a solution and the integrals representing the constants c_n fail to converge. The function $f(x) = x^{-1}$ furnishes a simple example. It is possible also to choose constants c_n [such as $(2n+1)P_n(u)P_n(v)$] for which the infinite series of Legendre functions converges for almost all positive values of x but the infinite series of terms $c_n U_n(t)$ fails to converge. It is doubtful, however, whether the integral equation has then a solution.

A simple example in which the present method of expansion leads to an exact solution of the integral equation is obtained by putting $c_n = z^n$. Then, if $|z| < 1$,

$$f(x) = \frac{1}{x+1} \sum_{n=0}^{\infty} z^n P_n\left(\frac{x-1}{x+1}\right) = [(x+1)^2(1+z^2) - 2z(x^2-1)]^{-\frac{1}{2}},$$

and so

$$g(t) = \frac{1}{1-z} \exp\left[-t\,\frac{1+z^2}{(1-z)^2}\right] I_0\left[\frac{2tz}{(1-z)^2}\right].$$

The polynomial $Z_n(t)$ thus has the generating function

$$\frac{1}{1-z} \exp\left[-\frac{2tz}{(1-z)^2}\right] I_0\left[\frac{2tz}{(1-z)^2}\right] = \sum_{n=0}^{\infty} z^n Z_n(t). \tag{15}$$

This may be verified with the aid of the equation[19]

$$(1-t)^{-m-1}(1+t)^m P_m\left(\frac{1+t^2}{1-t^2}\right) = \sum_{n=0}^{\infty} t^n F_n(-2m-1).$$

By expanding the generating function it is readily seen that

$$Z_n(t) = {}_2F_2(-n, n+1; 1, 1; t).$$

[19] H. Bateman, Tôhoku Math. Jour., vol. 37 (1933), p. 23.

This result is easily generalized. Indeed,

$$\begin{aligned}(16)\qquad &\int_0^\infty e^{-xt-t}t^b{}_2F_2(-n, a+n; c, b; t)dt\\ &= (1+x)^{-b-1}F(-n, a+n; c; u)\,\Gamma(b+1),\end{aligned}$$

where $u = (1 + x^{-1})$. Thus, expansions in series of Jacobi's polynomials may be used instead of expansions in series of Legendre polynomials. Since, if $C_n^\nu(z)$ is Gegenbauer's polynomial and $|z| < 1$,

$$\begin{aligned}(17)\qquad &\Gamma(1+b)(1+x)^{-b-1}C_n^\nu\left(\frac{x-1}{x+1}\right)\\ &= \frac{\Gamma(2\nu+n)}{n!\,\Gamma(2\nu)}\int_0^\infty e^{-xt-t}t^b{}_2F_2(-n, 2\nu+n; \nu+\tfrac{1}{2}, b; t)dt,\end{aligned}$$

we have

$$(18)\qquad \Gamma(1+b)(1+x)^{-b-1}\left(1 - 2z\frac{x-1}{x+1} + z^2\right)^{-\nu} = \int_0^\infty e^{-xt}g(t)dt,$$

where

$$e^t g(t) = t^b\sum_{n=0}^\infty \frac{\Gamma(2\nu+n)}{n!\,\Gamma(2\nu)} z^n{}_2F_2(-n, 2\nu+n; \nu+\tfrac{1}{2}, b; t).$$

3. A second way of using orthogonal functions for the expansion of $f(x)$ is to make use of the polynomials of Laguerre and the equation

$$(19)\qquad x^{-u-1}L_n^{v+\frac{1}{2}u}(x^{-1}) = \int_0^\infty e^{-xt}t^{\frac{1}{2}u}J(u, v, n; t^{\frac{1}{2}})dt,$$

where $x^{-u}J(u, v, n; x)$ denotes the polynomial[20]

$$\frac{\Gamma(v+\frac{1}{2}u+n+1)}{n!\,\Gamma(u+1)\Gamma(v+\frac{1}{2}u+1)}\,{}_1F_2(-n; u+1, v+\tfrac{1}{2}u+1; x^2),$$

which may be derived from the generating functions

$$(20)\qquad (1-t^2)^{-v-1}J_u[2xt(1-t^2)^{-\frac{1}{2}}] = \sum_{n=0}^\infty t^{2n+u}J(u, v, n; x) \qquad (|t| < 1),$$

$$\begin{aligned}(21)\qquad &e^t{}_0F_2(u+1, v+\tfrac{1}{2}u+1; -x^2t)\\ &= \Gamma(u+1)\Gamma(v+\tfrac{1}{2}u+1)\sum_{n=0}^\infty J(u, v, n; x)x^{-u}t^n/\Gamma(v+\tfrac{1}{2}u+n+1).\end{aligned}$$

These results may be verified by first using the well-known inversion formula ($C_{m,n}$ = binomial coefficient)

$$a_m = \sum_{n=0}^m C_{r,m-n}b_n, \qquad b_m = \sum_{n=0}^\infty C_{-r,m-n}a_n$$

[20] This polynomial may be used to construct a set of solutions of Laplace's equation in four independent variables.

to obtain the expansion

$$x^{u+2m} = m!\,\Gamma(u+m+1)\sum_{n=0}^{m}(-1)^n C_{m+\frac{1}{2}u+v,\,m-n}\,J(u,v,n;x), \tag{22}$$

from which the others are readily obtained by summation. Writing D_x for d/dx and $J_n^{u,v}$ for $J(u, v, n; x)$, we have the recurrence relations

$$\begin{gathered} D_x(x^u J_n^{u,v}) = 2x^u J_n^{u-1,\,v+\frac{1}{2}}, \qquad D_x(x^{-u} J_n^{u,v}) = -2x^{-u} J_{n-1}^{u+1,\,v+\frac{1}{2}}, \\ xJ_n^{u-1,v} + xJ_{n-1}^{u+1,v} = uJ_n^{u,v-\frac{1}{2}}, \qquad J_n^{u,v-1} = J_n^{u,v} - J_{n-1}^{u,v}, \\ D_x J_n^{u,v} = J_n^{u-1,\,v+\frac{1}{2}} - J_{n-1}^{u+1,\,v+\frac{1}{2}}, \end{gathered} \tag{23}$$

and the following relations are easily verified

$$J(u, v+m, n; x) = \sum_{s=0}^{\infty} C_{m+s-1,\,s}\,J(u, v, n-s; x), \tag{24}$$

$$\begin{gathered} x^{u-p}J_p(2ax) = \sum_{n=0}^{\infty}\Gamma(u+n+1)a^{p+2n}\cdot \\ F(u+n+1,\, n+\tfrac{1}{2}u+v+1,\, p+n+1;\, a^2)J(u,v,n;x)/\Gamma(p+n+1), \end{gathered} \tag{25}$$

$$\begin{gathered} \int_0^{\frac{\pi}{2}} J(u,v,n;x\sin\theta)(\sin\theta)^{u+1}(\cos\theta)^{2w-1}\,d\theta \\ = \tfrac{1}{2}\Gamma(w)x^{-w}J(u+w,\, v-\tfrac{1}{2}w,\, n; x), \qquad R(u) > 0, \qquad R(w) > 0, \end{gathered} \tag{26}$$

$$\begin{gathered} \int_0^{\infty} e^{-ax}J(u,v,n;x)x^u\,dx \\ = 2^{2u}a^{-2u-1}\Gamma(u+\tfrac{1}{2})C_{v+\frac{1}{2}u+n,\,n}\,\pi^{-\frac{1}{2}}F(-n,\, u+\tfrac{1}{2};\, v+\tfrac{1}{2}u+1;\, 4a^{-2}), \end{gathered} \tag{27}$$

$$\begin{gathered} \int_0^{\infty} K_s(ax)x^{u-s+1}J(u,v,n;x)\,dx \\ = 2^{2u-s}a^{s-2u-2}\Gamma(u-s+1)C_{v+\frac{1}{2}u+n,\,n}F(-n,\, u+1-s,\, v+\tfrac{1}{2}u+1;\, 4a^{-2}). \end{gathered} \tag{28}$$

The cases $u = -\frac{1}{2}$, $u = \frac{1}{2}$ are of special interest. We then write

$$(1-t^2)^{-v-1}\cos\,[2xt(1-t^2)^{-\frac{1}{2}}] = \sum_{n=0}^{\infty} t^{2n}\,Cl_n^v(x) \qquad (|t| < 1), \tag{29}$$

$$(1-t^2)^{-v-1}\sin\,[2xt(1-t^2)^{-\frac{1}{2}}] = \sum_{n=0}^{\infty} t^{2n+1}\,Sl_n^v(x) \qquad (|t| < 1), \tag{30}$$

$$D_x Cl_n^v(x) = -2Sl_{n-1}^{v+\frac{1}{2}}(x), \qquad D_x Sl_n^v(x) = 2Cl_n^{v+\frac{1}{2}}(x), \tag{31}$$

$$(2n+1)Sl_n^v(x) = 2x\,Cl_n^{v+\frac{1}{2}}(x) + (2v+2)Sl_{n-1}^{v+\frac{1}{2}}(x), \tag{32}$$

$$(2n+1)Sl_n^v(x) - (2n+1+2v)Sl_{n-1}^v(x) = xD_x Sl_n^v(x), \tag{33}$$

$$(n+1)Cl_{n+1}^v(x) = (v+1)Cl_n^{v+1}(x) - x\,Sl_n^{v+\frac{1}{2}}(x), \tag{34}$$

$$2n\,Cl_n^v(x) - (2n+2v)Cl_{n-1}^v(x) = xD_x Cl_n^v(x). \tag{35}$$

4. Instead of expanding $f(x)$ in a series of orthogonal functions associated with the range $(0, \infty)$ we may use a newtonian series. Writing

$$C_{\frac{1}{x}, n+1} = \int_0^\infty e^{-xt} N_n(t) dt,$$

where $N_n(t)$ is a polynomial of degree n in t, it is readily seen that, if $|z| < 1$,

$$zN_0(t) + z^2 N_1(t) + z^3 N_2(t) + \cdots = t^{-\frac{1}{2}}[\log(1+z)]^{\frac{1}{2}} I_1[2\{t \log(1+z)\}^{\frac{1}{2}}].$$

The polynomial $N_n(t)$ is closely related to Angelesco's polynomial[21] $A_n(x)$, which may be defined by means of the contour integral

$$A_n(x) = \frac{1}{2\pi i} \int_C \frac{dz}{z} e^z \left(\frac{x}{z} - 1\right)\left(\frac{x}{z} - 2\right) \cdots \left(\frac{x}{z} - n\right).$$

If

$$B_n(x) = (-1)^n \frac{d^{n+1}}{dx^{n+1}} [1 - e^{-x}]^{n+1},$$

Angelesco obtains the relation

$$\int_0^\infty A_m(x) B_n(x)\, dx = \delta_{mn},$$

which enables one to find the coefficients in a series of functions $A_m(x)$.

California Institute of Technology

[21] A. Angelesco, Jour. de Math., (9), vol. 2 (1923), p. 403. Our notation differs from that of Angelesco as we wish to avoid confusion between his polynomial, which he denotes by $P_n(x)$, and the Legendre polynomial. The contour in Angelesco's integral is a simple one enclosing the point $z = 0$. The function $A_n(x)$ is a solution of the integral equation

$$\frac{1}{x}\left(\frac{1}{x} - 1\right)\left(\frac{1}{x} - \frac{1}{2}\right) \cdots \left(\frac{1}{x} - \frac{1}{n}\right) = \int_0^\infty e^{-xt} A_n(t) dt.$$

ÜBER DEN FÜHRER EINES RINGES IN ALGEBRAISCHEN ZAHLKÖRPERN

VON MICHAEL BAUER

1. Es sei α eine primitive ganze Zahl des algebraischen Zahlkörpers $K = R(\alpha)$. Der Führer des Ringes $O(\alpha)$ ist bekanntlich ein Ideal, dasselbe soll durch $\mathfrak{t}(\alpha)$ bezeichnet werden. Nach Dedekind hat man

$$(1) \qquad \vartheta = \mathfrak{t}(\alpha)\mathfrak{d}, \qquad\qquad |D| = N(\mathfrak{d}),$$

wo ϑ, beziehungsweise $\mathfrak{d}$ die Differente der Zahl α, bzw. des Körpers K ist, D ist die Körperdiskriminante und N bedeutet die Norm.

Ore[1] hat den Führer des Ringes $O(\alpha)$ einer weiteren Untersuchung unterzogen, indem er den Führer inbezug auf die Primzahl p, bzw. auf das Primideal $\mathfrak{p}$ eingeführt hat. Die Ideale $\mathfrak{t}_p(\alpha)$ bzw. $\mathfrak{t}_\mathfrak{p}(\alpha)$ werden durch die ganzen Zahlen φ_1, bzw. φ_2 gebildet, für welche

$$\varphi_1\omega \equiv P_1^{(t)}(\alpha) \pmod{p^t} \text{ bzw. } \varphi_2\omega \equiv P_2^{(t)}(\alpha) \pmod{\mathfrak{p}^t}$$

ausfallen, wo $P_1^{(t)}$, $P_2^{(t)}$ rationale ganzzahlige Polynome sind. Die Zahl ω ist eine beliebige ganze Körperzahl, ferner ist t eine beliebige positive rationale ganze Zahl. Nun ist zunächst

$$(2) \qquad \mathfrak{t}(\alpha) = \prod_p \mathfrak{t}_p(\alpha).$$

Das Ideal $\mathfrak{t}_p(\alpha)$ lässt sich weiter zerlegen, genügt α der irreduziblen ganzzahligen Gleichung $F(x) = 0$, so spielt bei der Zerlegung von $\mathfrak{t}_p(\alpha)$ die Zerlegung von $F(x) \pmod{p^\nu}$ in irreduzible Faktoren eine entscheidende Rolle. Es sei im Körper K in Primideale zerlegt

$$(3) \qquad p = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_k^{e_k} \cdots \mathfrak{p}_r^{e_r},$$

wo $\mathfrak{p}_k$ ein Primideal g_k-ten Grades bezeichnet, dann ist in irreduzible Faktoren, deren höchste Koeffizienten gleich Eins sind, zerlegt

$$(3^*) \qquad F(x) \equiv F_1(x) \cdots F_k(x) \cdots F_r(x) \pmod{p^\nu}$$

wenn $\nu \geqq \delta + 1$ ausfällt, wo p^δ die höchste Potenz von p ist, die in der Diskriminante $D(F(x))$ enthalten ist. Gehört $\mathfrak{p}_k$ im Sinne von Ore[2] zum Polynom $F_k(x)$, so ist sein Grad $n_k = e_k g_k$. Es sei $F_l(\alpha)$ genau durch $\mathfrak{p}_k^{\gamma_{kl}}$ $(k \neq l)$ teilbar. Das Ideal $\mathfrak{t}_{\mathfrak{p}_k}(\alpha)$ ist eine Potenz von $\mathfrak{p}_k$, man kann setzen

$$(4) \qquad \mathfrak{t}_{\mathfrak{p}_k}(\alpha) = \mathfrak{p}_k^{\tau_k},$$

Received May 14, 1936.

[1] *Über den Zusammenhang zwischen den definierenden Gleichungen und der Idealtheorie in algebraischen Körpern*, Math. Annalen, Bd. 96 (1927), S. 313–352.

[2] A.a.O. S. 326.

und es ist nach Ore

$$(5) \qquad \mathfrak{t}_p(\alpha) = \prod_{k=1}^{r} \mathfrak{p}_k^{\tau_k+\gamma_k}, \qquad \gamma_k = \sum_{l=1}^{r} \gamma_{kl}, \qquad \gamma_{kk} = 0.$$

Die Zahlen γ_{kl} $(k \neq l)$ sind für $\nu \geqq \delta + 1$ invariant.

2. Die Relation (5) kann in zwei Teile getrennt werden. Zunächst wird behauptet

$$(6^*) \qquad \mathfrak{t}_p(\alpha) \text{ teilt das Ideal } \mathfrak{a}_p = \prod_{k=1}^{r} \mathfrak{p}_k^{\tau_k+\gamma_k}.$$

Diese Tatsache wird a.a.O. auf S. 336 bewiesen. Dann aber wird gezeigt, dass $\mathfrak{t}_p(\alpha)$ keinen echten Teiler von $\mathfrak{a}_p$ bildet, also besteht (5). Es kann aber leicht mit den von Ore benützten Mitteln, unabhängig von (6*) bewiesen werden,[3] dass

$$(6^{**}) \qquad \mathfrak{t}_p(\alpha) \text{ durch das Ideal } \mathfrak{a}_p \text{ teilbar ist.}$$

Es ist nicht ohne Interesse, dass die Relation (5) sowohl aus (6*), als auch aus (6**) gefolgert werden kann, wenn gewisse Sätze, welche von Ore in der genannten Abhandlung bewiesen wurden, angewendet werden. Diese Sätze werden im Folgenden zusammengestellt, der Beweis kann ohne die Theorie der Führerzerlegung geführt werden.[4]

Es seien p^{δ_k}, bzw. $p^{\rho_{kl}}$ $(k \neq l)$ die höchsten Potenzen von p, die in der Diskriminante $D(F_k(x))$, bzw. in der Resultante $R(F_k(x), F_l(x))$ enthalten sind. Es sei ferner

$$F_k(\bar{\alpha}_k) = 0$$

und die Indexe der Zahlen α bzw. $\bar{\alpha}_k$ bezüglich der Körper $K = R(\alpha)$ bzw. $K^{(k)} = R(\bar{\alpha}_k)$ sollen genau p^{ψ}, bzw. $p^{\bar{\psi}_k}$ enthalten, dann ist für genügend grosse[5] ν

$$(7) \qquad \psi = \sum_{k=1}^{r} \bar{\psi}_k + \sum_{k>l} \rho_{kl}$$

$$(8) \qquad \rho_{kl} = g_k \gamma_{kl} \qquad (k \neq l).$$

[3] Es ist zu beachten, dass im Falle $\gamma_k + \tau_k \geqq 1$, solche rat. ganzzahlige Polynome $\Pi_k(x)$, $G_k(x)$ vorhanden sind, wofür $\Pi_k(x)G_k(x) \not\equiv 0 \pmod{p}$ ausfällt und der Grad der linken Seite kleiner als der Körpergrad von K ist, ferner bildet $\Pi_k(\alpha)G_k(\alpha)p^{-1}$ eine ganze Zahl, welche genau durch $\mathfrak{p}_k^{\tau_k+\gamma_k-1}$ teilbar ist und eine beliebig hohe Potenz des Ideals $\mathfrak{p}_l$ $(l \neq k)$ enthält. Andererseits kann $\mathfrak{t}_p(\alpha)$ nur die Primideale $\mathfrak{p}_1, \cdots, \mathfrak{p}_k, \cdots, \mathfrak{p}_r$, als Teiler besitzen.

[4] Vgl. meine Note *Bemerkung zum Hensel-Oreschen Hauptsatze*, Acta Litt. ac Scient. reg. univ. Hungaricae Francisco-Josephinae, tom. 8. Wir können bei unseren Betrachtungen die Relationen (7), (9*), (9**) durch andere ersetzen, welche in Fricke, *Lehrb. der Algebra*, Bd. 3, Braunschweig, 1928, angegeben sind. Vgl. (12) S. 118 und (20) S. 114.

[5] Hier muss ν nicht möglichst klein gewählt werden.

Im Körper $K^{(k)} = R(\bar{\alpha}_k)$ wird $p = \bar{\mathfrak{p}}_k^{e_k}$, der Grad des Primideals $\bar{\mathfrak{p}}_k$ ist gleich g_k, aus der Definition des Führers folgt

$$(9^*) \qquad \mathfrak{t}_p(\bar{\alpha}_k) = \mathfrak{t}_{\bar{\mathfrak{p}}_k}(\bar{\alpha}_k)$$

und Ore hat bewiesen, dass

$$(9^{**}) \qquad \mathfrak{t}_{\bar{\mathfrak{p}}_k}(\bar{\alpha}_k) = \bar{\mathfrak{p}}_k^{\tau_k}$$

ist. Aus dem Vorhergehenden folgt

$$(10) \quad p^{2\psi} = N(\mathfrak{t}_p(\alpha)), \qquad 2\psi = \sum_{k=1}^{r} g_k\tau_k + \sum_{k=1}^{r}\sum_{l=1}^{r} g_k\gamma_{kl} = \sum_{k=1}^{r} g_k(\tau_k + \gamma_k).$$

Man kann aus (10) die Relation (5) beweisen, sobald irgendeine der Relationen (6*), bzw. (6**) feststeht. Es ist nämlich

$$\mathfrak{t}_p(\alpha) = \sum_{k=1}^{r} \mathfrak{p}_k^{\tau_k+\gamma_k+\epsilon_k}$$

wo $\epsilon_k \leqq 0$ bzw. $\epsilon_k \geqq 0$ ist, jenachdem (6*), oder (6**) gilt. Hieraus folgt

$$\sum_{k=1}^{r} g_k(\tau_k + \gamma_k + \epsilon_k) = \sum_{k=1}^{r} g_k(\tau_k + \gamma_k),$$

woraus sich $\epsilon_k = 0$ ergibt. Infolgedessen ist (5) richtig.

Budapest, Hungary.

DIRECT DECOMPOSITIONS

By Oystein Ore

One of the fundamental representations of algebraic systems is the decomposition into direct products. The principal theorem on direct decompositions has been proved for ideals in commutative and non-commutative rings under various assumptions; for groups it is the well-known *Schmidt-Remak* theorem. In a recent paper[1] *On the foundation of abstract algebra* I have shown that the main theorem on direct decompositions holds for all Dedekind structures which are of finite length or satisfy one chain condition and an additional restriction. It should be observed that such an additional condition is necessary, since it is known that the theorem is not true for all Dedekind structures satisfying one chain condition.[2]

In the present paper the properties of direct decompositions are studied further and various new and interesting facts about such decompositions are obtained. The method incidentally gives a new proof for the main theorem valid in the finite case or, with an additional restriction, when only one chain condition is satisfied. The method is a generalization of a method introduced by Krull[3] in the study of the so-called generalized abelian groups. The results of Krull are extended and refined in various ways and the theory is greatly simplified by the use of structures. It may not be superfluous to observe again that the formulation of the theory in terms of structures gives it a great generality, making it valid for all algebraic systems in which the Dedekind axiom is satisfied. As an example let us mention that the present theory is valid for arbitrary groups, a case to which the original theory of Krull was not applicable.

Chapter 1. The main theorem

1. Theorems on components. Before we can begin the principal investigations it is necessary to mention a few facts about the so-called *components*. The components have already been defined in II, Chap. 1, but we shall have to recall some of their properties here. Let

$$\mathfrak{M} = [\mathfrak{B}_1, \mathfrak{B}_2], \qquad (\mathfrak{B}_1, \mathfrak{B}_2) = \mathfrak{E}_0 \tag{1}$$

Received April 13, 1936.

[1] *On the foundation of abstract algebra*, I, II, Annals of Math., vol. 36 (1935), pp. 406–437, vol. 37 (1936), pp. 265–292. These articles will be referred to in the following as Ore I and Ore II and the terminology of these papers will be used in the following without further explanation.

[2] E. Steinitz, *Rechteckige Systeme und Moduln in algebraischen Zahlkörpern*, I, Math. Ann., vol. 71 (1911), pp. 328–354. W. Krull, *Matrizen, Moduln und verallgemeinerte Abelsche Gruppen*, Sitzungsberichte Heidelberg, 1932, pp. 13–38. See also the discussion in Ore II.

[3] W. Krull, *Über verallgemeinerte endliche Abelsche Gruppen*, Math. Zeitschr., vol. 23 (1925), pp. 161–196.

be a direct decomposition of a quotient $\mathfrak{M}$ in a Dedekind structure Σ. Here $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are quotients with the same denominator and $\mathfrak{E}_0$ is the corresponding unit quotient. Now let $\mathfrak{S}$ be a factor of $\mathfrak{M}$. The quotients

$$\mathfrak{B}_1\{\mathfrak{S}\} = (\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{S}]), \qquad \mathfrak{B}_2\{\mathfrak{S}\} = (\mathfrak{B}_2, [\mathfrak{B}_1, \mathfrak{S}]) \tag{2}$$

are then called the *components*[4] of $\mathfrak{S}$ in $\mathfrak{B}_1$ and $\mathfrak{B}_2$.

One finds that the union of the components of two factors $\mathfrak{S}_1$ and $\mathfrak{S}_2$ in the same quotient $\mathfrak{B}_i$ is equal to the component of the union $[\mathfrak{S}_1, \mathfrak{S}_2]$ in the same $\mathfrak{B}_i$:

$$\mathfrak{B}_i\{[\mathfrak{S}_1, \mathfrak{S}_2]\} = [\mathfrak{B}_i\{\mathfrak{S}_1\}, \mathfrak{B}_i\{\mathfrak{S}_2\}]. \tag{3}$$

We may now prove

THEOREM 1. *Let $\mathfrak{S}_1 \geqq \mathfrak{S}_2$ be two factors of $\mathfrak{M}$. Then there exists the following similarity relation between their components in $\mathfrak{B}_1$:*

$$\mathfrak{B}_1\{\mathfrak{S}_1\}/\mathfrak{B}_1\{\mathfrak{S}_2\} \cong \mathfrak{S}_1/[\mathfrak{S}_2, (\mathfrak{B}_2, \mathfrak{S}_1)]. \tag{4}$$

One finds by transformation with $\mathfrak{B}_2$ that the last quotient is similar to $[\mathfrak{B}_2, \mathfrak{S}_1]/[\mathfrak{B}_2, \mathfrak{S}_2]$ and a contraction with $\mathfrak{B}_1$ gives the desired similarity (4). A consequence of Theorem 1 is the

LEMMA. *The component $\mathfrak{B}_1\{\mathfrak{S}\}$ is similar to a left-hand factor of $\mathfrak{S}$:*

$$\mathfrak{B}_1\{\mathfrak{S}\} \cong \mathfrak{S}/(\mathfrak{S}, \mathfrak{B}_2). \tag{5}$$

From this lemma it follows that $\mathfrak{B}_1\{\mathfrak{S}\}$ is the unit quotient if and only if $\mathfrak{S}$ is contained in $\mathfrak{B}_2$ and $\mathfrak{B}_1\{\mathfrak{S}\}$ is similar to $\mathfrak{S}$ if $\mathfrak{S}$ is relatively prime to $\mathfrak{B}_2$.

THEOREM 2. *Let $\mathfrak{C}$ be a factor of $\mathfrak{M}$ and let $\mathfrak{S}_1$ and $\mathfrak{S}_2$ be two factors of $\mathfrak{C}$. The necessary and sufficient condition that $\mathfrak{S}_1$ and $\mathfrak{S}_2$ have the same component in $\mathfrak{C}_1$ is that*

$$[\mathfrak{S}_1, (\mathfrak{B}_2, \mathfrak{C})] = [\mathfrak{S}_2, (\mathfrak{B}_2, \mathfrak{C})]. \tag{6}$$

From the identity

$$(\mathfrak{B}_1, [\mathfrak{S}_1, \mathfrak{B}_2]) = (\mathfrak{B}_1, [\mathfrak{S}_2, \mathfrak{B}_2])$$

one obtains by taking the union of both sides with $\mathfrak{B}_2$ and applying the Dedekind axiom

$$[\mathfrak{S}_1, \mathfrak{B}_2] = [\mathfrak{S}_2, \mathfrak{B}_2],$$

and (6) follows by taking the cross-cut of both sides with $\mathfrak{C}$. The sufficiency of the condition (6) follows directly from (3). As a special case of Theorem 2

[4] These components are easily seen to correspond to the ordinary notion of components, for instance in group theory, where the component of a sub-group $\mathfrak{S}$ in $\mathfrak{A}$ is the subgroup of $\mathfrak{A}$ consisting of all elements α occurring in the direct product representations $\sigma = \alpha\beta$ of the elements σ of $\mathfrak{S}$.

we see that the necessary and sufficient condition that any two factors $\mathfrak{S}_1$ and $\mathfrak{S}_2$ of $\mathfrak{M}$ have the same components in $\mathfrak{B}_1$ is

$$(7) \qquad [\mathfrak{S}_1, \mathfrak{B}_2] = [\mathfrak{S}_2, \mathfrak{B}_2].$$

THEOREM 3. *Let $\mathfrak{C}$ be a factor of $\mathfrak{M}$ and $\mathfrak{B}_1'$ some factor of $\mathfrak{B}_1$. The maximal factor of $\mathfrak{B}_1'$ which is the $\mathfrak{B}_1$-component of some factor $\mathfrak{C}_1$ of $\mathfrak{C}$ is*

$$\mathfrak{B}_1'' = (\mathfrak{B}_1', \mathfrak{B}_1\{\mathfrak{C}\}),$$

and the maximal factor of $\mathfrak{C}$ having its $\mathfrak{B}_1$-component equal to $\mathfrak{B}_1''$ is

$$(8) \qquad \mathfrak{C}_1 = (\mathfrak{C}, [\mathfrak{B}_1', \mathfrak{B}_2]).$$

Since any factor of $\mathfrak{C}$ has a $\mathfrak{B}_1$-component contained in $\mathfrak{B}_1\{\mathfrak{C}\}$, it follows that $\mathfrak{B}_1''$ is the maximal factor of $\mathfrak{B}_1'$, which may be the $\mathfrak{B}_1$-component of some factor of $\mathfrak{C}$. It is easily verified by means of the Dedekind axiom that the quotient $\mathfrak{C}_1$ defined in (8) actually has the $\mathfrak{B}_1$-component $\mathfrak{B}_1''$. Furthermore it follows from Theorem 2 that $\mathfrak{C}_1$ is the maximal such factor, since it contains $(\mathfrak{B}_2, \mathfrak{C})$. As a consequence of Theorem 3 we find that the necessary and sufficient condition that a factor $\mathfrak{B}_1'$ be the $\mathfrak{B}_1$-component of some factor of $\mathfrak{C}$ is that $\mathfrak{B}_1'$ be contained in $\mathfrak{B}_1\{\mathfrak{C}\}$.

These results have all been derived for right-hand divisibility, but the analogous results hold for left-hand divisibility.

2. **Reductions and increments.** Let us now suppose that there exist two direct decompositions of the same quotient

$$(9) \qquad \mathfrak{M} = [\mathfrak{A}_1, \mathfrak{A}_2] = [\mathfrak{B}_1, \mathfrak{B}_2], \qquad (\mathfrak{A}_1, \mathfrak{A}_2) = (\mathfrak{B}_1, \mathfrak{B}_2) = \mathfrak{E}_0.$$

We shall now introduce a new operation which consists in the repetition of the operation of taking components. Let $\mathfrak{S}$ be a factor of $\mathfrak{A}_1$ and $\mathfrak{B}_1\{\mathfrak{S}\}$ its component in $\mathfrak{B}_1$. This component again has a component in $\mathfrak{A}_1$, namely

$$(10) \qquad \mathfrak{R}_1^{(1)}\{\mathfrak{S}\} = \mathfrak{A}_1\{\mathfrak{B}_1\{\mathfrak{S}\}\} = (\mathfrak{A}_1, [\mathfrak{A}_2, (\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{A}_1])])$$

and this quotient shall be called the *reduction* of $\mathfrak{S}$ (with respect to $\mathfrak{B}_1$). For the special case $\mathfrak{S} = \mathfrak{A}_1$ we shall write

$$(11) \qquad \mathfrak{R}_1^{(1)} = \mathfrak{R}_1^{(1)}\{\mathfrak{A}_1\}.$$

A second fundamental operation is derived from the consideration of the following problem: *For a given factor $\mathfrak{S}$ of $\mathfrak{A}_1$ we wish to determine the maximal factor of $\mathfrak{A}_1$ such that its reduction is contained in $\mathfrak{S}$.* Through a double application of Theorem 3 we find that the quotient in question must be

$$(12) \qquad N_1^{(1)}\{\mathfrak{S}\} = (\mathfrak{A}_1, [\mathfrak{B}_2, (\mathfrak{B}_1, [\mathfrak{A}_2, \mathfrak{S}])]),$$

and we shall call it the increment of $\mathfrak{S}$ (with respect to $\mathfrak{B}_1$). The increment of the unit factor $\mathfrak{E}_0$ obviously represents the maximal factor of $\mathfrak{A}_1$ having its reduction equal to $\mathfrak{E}_0$. This factor

$$(13) \qquad \mathfrak{N}_1^{(1)} = N_1^{(1)}\{\mathfrak{E}_0\} = (\mathfrak{A}_1, [\mathfrak{B}_2, (\mathfrak{B}_1, \mathfrak{A}_2)])$$

we shall call the *null-factor* of $\mathfrak{A}_1$ (with respect to $\mathfrak{B}_1$).

Before we study the properties of these new concepts, let us for a moment consider the corresponding left-hand concepts. These furnish us with a dualism between reduction and increment which manifests itself in all the following investigations.

To the given direct decompositions (9) there exist corresponding left-hand decompositions

$$\mathfrak{M} = [\mathfrak{A}_1^*, \mathfrak{A}_2^*]_l = [\mathfrak{B}_1^*, \mathfrak{B}_2^*]_l, \tag{14}$$

where

$$\mathfrak{A}_1^* = \mathfrak{M}/\mathfrak{A}_2, \qquad \mathfrak{A}_2^* = \mathfrak{M}/\mathfrak{A}_1, \qquad \mathfrak{B}_1^* = \mathfrak{M}/\mathfrak{B}_2, \qquad \mathfrak{B}_2^* = \mathfrak{M}/\mathfrak{B}_1,$$

are quotients similar to $\mathfrak{A}_1$, $\mathfrak{A}_2$, $\mathfrak{B}_1$, $\mathfrak{B}_2$ respectively. There exists a structure isomorphism between the r.h. factors $\mathfrak{S}$ of $\mathfrak{A}_1$ and the l.h. factors $\mathfrak{S}^*$ of $\mathfrak{A}_1^*$ given by the correspondence

$$\mathfrak{S}^* \to \mathfrak{M}/[\mathfrak{A}_2, \mathfrak{S}]. \tag{15}$$

The l.h. reduction and increment of $\mathfrak{S}^*$ with respect to $\mathfrak{B}_1^*$ are then easily found to be

$$\begin{aligned} R_1^*\{\mathfrak{S}^*\} &= \mathfrak{M}/[\mathfrak{A}_2, N_1^{(1)}\{\mathfrak{S}\}], \\ N_1^*\{\mathfrak{S}^*\} &= \mathfrak{M}/[\mathfrak{A}_2, R_1^{(1)}\{\mathfrak{S}\}], \end{aligned} \tag{16}$$

and hence the correspondence (15) makes l.h. reduction correspond to r.h. increment and l.h. increment to r.h. reduction. From this correspondence one also obtains the following similarity relations

$$R_1^*\{\mathfrak{S}^*\} \cong N_1^{(1)}\{\mathfrak{S}\}, \qquad N_1^*\{\mathfrak{S}\} \cong R_1^{(1)}\{\mathfrak{S}\},$$

and hence one obtains for $\mathfrak{S} = \mathfrak{E}_0$ and $\mathfrak{S} = \mathfrak{A}_i$

$$R_1^* \cong \mathfrak{N}_1^{(1)}, \qquad \mathfrak{N}_1^* \cong R_1^{(1)}.$$

3. **Theorems on reductions and increments.** We shall now derive a series of properties of these operations. From the relation (3) and the definition, the reduction must have the distributive property

$$R_1^{(1)}\{[\mathfrak{S}_1, \mathfrak{S}_2]\} = [R_1^{(1)}\{\mathfrak{S}_1\}, R_1^{(1)}\{\mathfrak{S}_2\}], \tag{17}$$

and in a similar way one proves

$$N_1^{(1)}\{(\mathfrak{S}_1, \mathfrak{S}_2)\} = (N_1^{(1)}\{\mathfrak{S}_1\}, N_1^{(1)}\{\mathfrak{S}_2\}). \tag{18}$$

From these relations we obtain the fact that if $\mathfrak{S}_1 \geqq \mathfrak{S}_2$ are two factors of $\mathfrak{A}_1$, then

$$R_1^{(1)}\{\mathfrak{S}_1\} \geqq R_1^{(1)}\{\mathfrak{S}_2\}, \qquad N_1^{(1)}\{\mathfrak{S}_1\} \geqq N_1^{(1)}\{\mathfrak{S}_2\}.$$

Very important are the following relations, which may be proved from the definitions through a repeated application of the Dedekind axiom.

THEOREM 4. *Let $\mathfrak{S}$ and $\mathfrak{T}$ be the two factors of $\mathfrak{A}_1$. Then the following identities hold*:

$$(19)\qquad \begin{aligned} N_1^{(1)}\{[R_1^{(1)}\{\mathfrak{S}\}, \mathfrak{T}]\} &= [\mathfrak{S}, N_1^{(1)}\{\mathfrak{T}\}], \\ R_1^{(1)}\{(N_1^{(1)}\{\mathfrak{S}\}, \mathfrak{T})\} &= (\mathfrak{S}, R_1^{(1)}\{\mathfrak{T}\}). \end{aligned}$$

For $\mathfrak{T} = \mathfrak{E}_0$ and $\mathfrak{T} = \mathfrak{A}_1$ we obtain the simpler identities

$$(20)\qquad \begin{aligned} N_1^{(1)}\{R_1^{(1)}\{\mathfrak{S}\}\} &= [\mathfrak{S}, \mathfrak{R}_1^{(1)}], \\ R_1^{(1)}\{N_1^{(1)}\{\mathfrak{S}\}\} &= (\mathfrak{S}, \mathfrak{R}_1^{(1)}), \end{aligned}$$

and hence also

$$(21)\qquad N_1^{(1)}\{\mathfrak{R}_1^{(1)}\} = \mathfrak{A}_1, \qquad R_1^{(1)}\{\mathfrak{R}_1^{(1)}\} = \mathfrak{E}_0.$$

Let us also mention the following symbolic identities, which one can derive from (20):

$$N \cdot R \cdot N = N, \qquad R \cdot N \cdot R = R.$$

By means of these relations we prove easily

THEOREM 5. *The necessary and sufficient condition that two factors $\mathfrak{S}_1$ and $\mathfrak{S}_2$ of $\mathfrak{A}_1$ have the same reductions is*

$$(22)\qquad [\mathfrak{S}_1, \mathfrak{R}_1^{(1)}] = [\mathfrak{S}_2, \mathfrak{R}_1^{(1)}].$$

The necessary and sufficient condition that they have the same increments is

$$(23)\qquad (\mathfrak{S}_1, \mathfrak{R}_1^{(1)}) = (\mathfrak{S}_2, \mathfrak{R}_1^{(1)}).$$

To prove that

$$(24)\qquad R_1^{(1)}\{\mathfrak{S}_1\} = R_1^{(1)}\{\mathfrak{S}_2\}$$

if and only if (22) is satisfied, we observe that the relation (22) follows from (24) by applying (20), and (24) follows from (22) by means of (17). The condition (23) is proved in a similar way.

THEOREM 6. *Let $\mathfrak{S}_1 \geqq \mathfrak{S}_2$ be two factors of $\mathfrak{A}_1$. Then the following similarity relations hold*:

$$(25)\qquad \begin{aligned} R_1^{(1)}\{\mathfrak{S}_1\}/R_1^{(1)}\{\mathfrak{S}_2\} &\cong \mathfrak{S}_1/[\mathfrak{S}_2, (\mathfrak{S}_1, \mathfrak{R}_1^{(1)})], \\ N_1^{(1)}\{\mathfrak{S}_1\}/N_1^{(1)}\{\mathfrak{S}_2\} &\cong (\mathfrak{S}_1, [\mathfrak{S}_2, \mathfrak{R}_1^{(1)}])/\mathfrak{S}_2, \end{aligned}$$

and for $\mathfrak{S}_1 = \mathfrak{A}_1$, $\mathfrak{S}_2 = \mathfrak{E}_0$ follows from either relation

$$\mathfrak{A}_i/\mathfrak{R}_1^{(1)} \cong \mathfrak{R}_1^{(1)}.$$

The first similarity relation (25) may be obtained by applying Theorem 1 twice, and the second follows dualistically.

Let us finally mention

Theorem 7. *Between the reductions and null-factors of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$ and $\mathfrak{B}_2$ there exist the relations*

$$\mathfrak{A}_1 = [\mathfrak{R}_1^{(1)}, \mathfrak{R}_2^{(1)}], \qquad \mathfrak{E}_0 = (\mathfrak{N}_1^{(1)}, \mathfrak{N}_2^{(1)}). \tag{26}$$

According to (3) we have for any factor $\mathfrak{S}$ of $\mathfrak{A}_1$

$$[R_1^{(1)}\{\mathfrak{S}\}, R_2^{(1)}\{\mathfrak{S}\}] = (\mathfrak{A}_1, [\mathfrak{A}_2, \mathfrak{B}_1', \mathfrak{B}_2']),$$

where $\mathfrak{B}_1'$ and $\mathfrak{B}_2'$ are components of $\mathfrak{S}$ in $\mathfrak{B}_1$ and $\mathfrak{B}_2$ respectively. Since

$$[\mathfrak{B}_1', \mathfrak{B}_2'] = [(\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{S}]), (\mathfrak{B}_2, [\mathfrak{B}_1, \mathfrak{S}])] = ([\mathfrak{B}_1, \mathfrak{S}]\,[\mathfrak{B}_2, \mathfrak{S}]) > \mathfrak{S},$$

we also find

$$[R_1^{(1)}\{\mathfrak{S}\}, R_2^{(1)}\{\mathfrak{S}\}] > \mathfrak{S},$$

a relation which gives the first of (26) as a special case. The second relation (26) may be obtained directly from the definition of $\mathfrak{N}_1^{(1)}$ and $\mathfrak{N}_2^{(1)}$.

4. **The algorithms.** We shall now consider the algorithms consisting in repeated application of the two operations of taking reductions and increments. We define the *n-th reduction* $R_1^{(n)}\{\mathfrak{S}\}$ of a factor $\mathfrak{S}$ of $\mathfrak{A}_1$ as the result of taking n successive reductions of $\mathfrak{S}$. We define similarly the *n-th increment* $N_1^{(n)}\{\mathfrak{S}\}$. Let us also write

$$\mathfrak{R}_1^{(n)} = R_1^{(n)}\{\mathfrak{A}_1\}, \qquad \mathfrak{N}_1^{(n)} = N_1^{(n)}\{\mathfrak{E}_0\},$$

where $\mathfrak{N}_1^{(n)}$ shall be called the *n-th null-factor* of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$. It is the maximal factor of $\mathfrak{A}_1$ for which the n-th reduction is equal to the unit quotient. The relations between left-hand and right-hand n-th reductions and increments are the same as those indicated in (16) for the case $n = 1$, reductions corresponding to increments and increments corresponding to reductions through the isomorphism defined by (15).

One may now derive properties for these general operations corresponding closely to those formerly obtained in the case $n = 1$. We observe first that for $R_1^{(n)}\{\mathfrak{S}\}$ and $I_1^{(n)}\{\mathfrak{S}\}$ we have the distributive properties expressed by (17) and (18). Through induction the relations (19) may be generalized to

$$\begin{aligned} N_1^{(n)}\{[R_1^{(n)}\{\mathfrak{S}\}, \mathfrak{T}]\} &= [\mathfrak{S}, N_1^{(n)}\{\mathfrak{T}\}], \\ R_1^{(n)}\{(N_1^{(n)}\{\mathfrak{S}\}, \mathfrak{T})\} &= (\mathfrak{S}, R_1^{(n)}\{\mathfrak{T}\}), \end{aligned} \tag{27}$$

and for $\mathfrak{T} = \mathfrak{E}_0$ and $\mathfrak{T} = \mathfrak{A}_i$ we obtain the special cases

$$\begin{aligned} N_1^{(n)}\{R_1^{(n)}\{\mathfrak{S}\}\} &= [\mathfrak{S}, \mathfrak{N}_1^{(n)}], \\ R_1^{(n)}\{N_1^{(n)}\{\mathfrak{S}\}\} &= (\mathfrak{S}, \mathfrak{R}_1^{(n)}). \end{aligned} \tag{28}$$

From these relations we obtain the following generalization of Theorem 5:

Theorem 8. *The necessary and sufficient condition that two factors $\mathfrak{S}_1$ and $\mathfrak{S}_2$ have the same n-th reduction is*

$$[\mathfrak{S}_1, \mathfrak{N}_1^{(n)}] = [\mathfrak{S}_2, \mathfrak{N}_1^{(n)}]. \tag{29}$$

The necessary and sufficient condition that they have the same n-th increment is

$$(\mathfrak{S}_1, \mathfrak{N}_1^{(n)}) = (\mathfrak{S}_2, \mathfrak{N}_1^{(n)}). \tag{30}$$

The proof is analagous to that of Theorem 5. From Theorem 6 one obtains by induction the more general similarity relations valid for $\mathfrak{S}_1 \geqq \mathfrak{S}_2$:

$$\begin{aligned} R_1^{(n)}\{\mathfrak{S}_1\}/R_1^{(n)}\{\mathfrak{S}_2\} &\cong \mathfrak{S}_1/[\mathfrak{S}_2, (\mathfrak{S}_1, \mathfrak{N}_1^{(n)})], \\ N_1^{(n)}\{\mathfrak{S}_1\}/N_1^{(n)}\{\mathfrak{S}_2\} &\cong (\mathfrak{S}_1, [\mathfrak{S}_2, \mathfrak{N}_1^{(n)}])/\mathfrak{S}_2, \end{aligned} \tag{31}$$

and again for $\mathfrak{S}_1 = \mathfrak{A}_1$, $\mathfrak{S}_2 = \mathfrak{E}_0$,

$$\mathfrak{A}_1/\mathfrak{N}_1^{(n)} \cong \mathfrak{R}_1^{(n)}. \tag{32}$$

Let us finally prove

Theorem 9. *For all m and n we have*

$$\mathfrak{A}_1 = [\mathfrak{R}_1^{(n)}, \mathfrak{R}_2^{(m)}], \tag{33}$$

and

$$\mathfrak{E}_0 = (\mathfrak{N}_1^{(n)}, \mathfrak{N}_2^{(m)}). \tag{34}$$

This theorem is true for $n = m = 1$ according to Theorem 5, and hence it may be proved by induction. We shall suppose that the theorem holds for all $n + m < N_0$ and prove that it is also true when $n + m = N_0$.

Let us observe that

$$\mathfrak{R}_1^{(n-1)} \geqq \mathfrak{R}_1^{(n)}, \qquad \mathfrak{R}_2^{(m-1)} \geqq \mathfrak{R}_2^{(m)}.$$

It follows then by the Dedekind axiom from the induction condition that

$$\mathfrak{A}_1 = ([\mathfrak{R}_1^{(n)}, \mathfrak{R}_2^{(m-1)}], [\mathfrak{R}_1^{(n-1)}, \mathfrak{R}_2^{(m)}]) = [\mathfrak{R}_1^{(n)}, \mathfrak{R}_2^{(m)}, (\mathfrak{R}_1^{(n-1)}, \mathfrak{R}_2^{(m-1)})],$$

and hence it is sufficient to show that

$$\mathfrak{U} = [\mathfrak{R}_1^{(n)}, \mathfrak{R}_2^{(m)}] \geqq (\mathfrak{R}_1^{(n-1)}, \mathfrak{R}_2^{(m-1)}). \tag{35}$$

However, since $\mathfrak{U}$ is the union of the $\mathfrak{A}_1$-components of $\mathfrak{B}_1\{\mathfrak{R}_1^{(n-1)}\}$ and $\mathfrak{B}_2\{\mathfrak{R}_2^{(m-1)}\}$, it is also according to (3) the $\mathfrak{A}_1$-component of

$$\begin{aligned} [\mathfrak{B}_1\{\mathfrak{R}_1^{(n-1)}\}, \mathfrak{B}_2\{\mathfrak{R}_2^{(m-1)}\}] &= [(\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{R}_1^{(n-1)}]), (\mathfrak{B}_2, [\mathfrak{B}_1, \mathfrak{R}_2^{(m-1)}])] \\ &= ([\mathfrak{B}_2, \mathfrak{R}_1^{(n-1)}], [\mathfrak{B}_1, \mathfrak{R}_2^{(m-1)}]) > (\mathfrak{R}_1^{(n-1)}, \mathfrak{R}_2^{(m-1)}), \end{aligned}$$

and the relation (35) follows immediately.

The relation (34) follows by a dualistic process. We have

$$\mathfrak{E}_0 = [(\mathfrak{N}_1^{(n)}, \mathfrak{N}_2^{(m-1)}), (\mathfrak{N}_1^{(n-1)}, \mathfrak{N}_2^{(m)})] = (\mathfrak{N}_1^{(n)}, \mathfrak{N}_2^{(m)}, [\mathfrak{N}_1^{(n-1)}, \mathfrak{N}_2^{(m-1)}]),$$

and hence it is sufficient to prove that

$$\mathfrak{V} = (\mathfrak{N}_1^{(n)}, \mathfrak{N}_1^{(m)}) \leqq [\mathfrak{N}_1^{(n-1)}, \mathfrak{N}_2^{(m-1)}]. \tag{36}$$

Through simple reductions one finds

$$\mathfrak{V} = (\mathfrak{A}_1, [(\mathfrak{B}_1, [\mathfrak{N}_1^{(n-1)}, \mathfrak{A}_2]), (\mathfrak{B}_2, [\mathfrak{N}_2^{(m-1)}, \mathfrak{A}_2])]) \leqq (\mathfrak{A}_1, [\mathfrak{A}_2, \mathfrak{N}_1^{(n-1)}, \mathfrak{N}_2^{(m-1)}]),$$

and from this relation (36) follows immediately.

5. **The invariants.** It follows from their definition that the reductions form a decreasing sequence

$$\mathfrak{A}_1 \geqq \mathfrak{R}_1^{(1)} \geqq \mathfrak{R}_1^{(2)} \geqq \cdots, \tag{37}$$

while the null-factors form an increasing sequence

$$\mathfrak{E}_0 \leqq \mathfrak{N}_1^{(1)} \leqq \mathfrak{N}_1^{(2)} \leqq \cdots. \tag{38}$$

If at any point of these sequences the equality sign holds, it must hold for all following terms. Let us now suppose that in the sequence (37) all terms become equal after the n-th. We shall then write

$$\mathfrak{R}_{1,1} = \mathfrak{R}_1^{(n)} = \mathfrak{R}_1^{(n+1)} = \cdots, \tag{39}$$

and we shall call $\mathfrak{R}_{1,1}$ the *reduction invariant* of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$. The quotient $\mathfrak{R}_{1,1}$ must always exist when the descending chain condition is satisfied in $\mathfrak{M}$ (or in $\mathfrak{A}_1$). The name is justified by the fact that

$$R_1\{\mathfrak{R}_{1,1}\} = \mathfrak{R}_{1,1}. \tag{40}$$

Similarly, if the terms in the sequence (38) all become equal after the m-th, we write

$$\mathfrak{N}_{1,1} = \mathfrak{N}_1^{(m)} = \mathfrak{N}_1^{(m+1)} = \cdots, \tag{41}$$

and we call $\mathfrak{N}_{1,1}$ the *increment invariant* of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$. It has the property

$$N_1\{\mathfrak{N}_{1,1}\} = \mathfrak{N}_{1,1}, \tag{42}$$

and it must always exist when the ascending chain condition is satisfied in $\mathfrak{M}$.

We can now prove

THEOREM 10. *When the invariant $\mathfrak{R}_{1,1}$ exists and is defined by* (39), *then*

$$\mathfrak{A}_1 = [\mathfrak{R}_{1,1}, \mathfrak{R}_1^{(n)}], \tag{43}$$

and if $\mathfrak{N}_{1,1}$ exists and is defined by (41), *then*

$$\mathfrak{E}_0 = (\mathfrak{N}_{1,1}, \mathfrak{N}_1^{(m)}). \tag{44}$$

Theorem 10 is a consequence of Theorem 8, since when $\mathfrak{R}_{1,1}$ exists, the quotients $\mathfrak{A}_1$ and $\mathfrak{R}_{1,1}$ have the same n-th reduction. The relation (44) follows in a similar way. It may be observed that according to (39) and (41) the relations (43) and (44) will also hold for all larger n and m.

From Theorem 10 we obtain in turn

THEOREM 11. *If both invariants $\mathfrak{R}_{1,1}$ and $\mathfrak{N}_{1,1}$ exist, $\mathfrak{A}_1$ has a direct decomposition*

$$\mathfrak{A} = [\mathfrak{R}_{1,1}, \mathfrak{N}_{1,1}], \qquad (\mathfrak{R}_{1,1}, \mathfrak{N}_{1,1}) = \mathfrak{E}_0. \tag{45}$$

We shall say that $\mathfrak{A}_1$ has a *regular decomposition* with respect to $\mathfrak{B}_1$, when the relation (45) holds. It is obvious that a regular decomposition always exists when $\mathfrak{M}$ (or $\mathfrak{A}_1$ only) has a finite length. We shall see later that it also holds

under much more general conditions. One very interesting property of the regular decomposition (45) is that it is explicitly expressible by means of the components in the given decomposition (9).

6. **Properties of invariants.** We shall now only consider quotients $\mathfrak{M}$ having the property that the regular decompositions exist for any component in any direct decomposition of $\mathfrak{M}$. Such a quotient $\mathfrak{M}$ or structure Σ may be called a *regular* quotient or structure. We shall discuss later the condition for a quotient or structure to have this property.

For such regular quotients we can prove

THEOREM 12. *Let $\mathfrak{R}_{1,1}$ be the reduction invariant and $\mathfrak{N}_{1,1}$ the increment invariant of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$, and similarly $\bar{\mathfrak{R}}_{1,1}$ and $\bar{\mathfrak{N}}_{1,1}$ the corresponding quotients for $\mathfrak{B}_1$ with respect to $\mathfrak{A}_1$. Then $\bar{\mathfrak{R}}_{1,1}$ is the $\mathfrak{B}_1$-component of $\mathfrak{R}_{1,1}$ and $\mathfrak{R}_{1,1}$ is the $\mathfrak{A}_1$-component of $\bar{\mathfrak{R}}_{1,1}$,*

$$(46) \qquad \bar{\mathfrak{R}}_{1,1} = (\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{R}_{1,1}]), \qquad \mathfrak{R}_{1,1} = (\mathfrak{A}_1, [\mathfrak{A}_2, \bar{\mathfrak{R}}_{1,1}]).$$

Similarly, one finds

$$(47) \qquad \bar{\mathfrak{N}}_{1,1} = (\mathfrak{B}_1, [\mathfrak{A}_2, \mathfrak{N}_{1,1}]), \qquad \mathfrak{N}_{1,1} = (\mathfrak{A}_1, [\mathfrak{B}_2, \bar{\mathfrak{N}}_{1,1}]).$$

It follows from the definition of the invariants that the $\mathfrak{B}_1$-component of $\mathfrak{R}_{1,1}$ must be unchanged when one takes reductions of it in $\mathfrak{A}_1$, and hence we have

$$\bar{\mathfrak{R}}_{1,1} \geqq (\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{R}_{1,1}]).$$

In the same way one finds

$$\mathfrak{R}_{1,1} \geqq (\mathfrak{A}_1, [\mathfrak{A}_2, \bar{\mathfrak{R}}_{1,1}]),$$

and the substitution of one relation into the other shows that the equality sign must hold. The relations (47) are proved in a similar manner.

THEOREM 13. *When $\mathfrak{R}_{1,1}$ and $\mathfrak{R}_{1,2}$ are the reduction invariants of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$ and $\mathfrak{B}_2$, similarly, $\mathfrak{N}_{1,1}$ and $\mathfrak{N}_{1,2}$ the corresponding increment invariants, then*

$$(48) \qquad \mathfrak{A}_1 = [\mathfrak{R}_{1,1}, \mathfrak{R}_{1,2}], \qquad \mathfrak{E}_0 = (\mathfrak{N}_{1,1}, \mathfrak{N}_{1,2}).$$

This theorem is a consequence of Theorem 9. We mention also the following result:

THEOREM 14. *The two equivalent conditions*

$$\mathfrak{R}_{1,1} = \mathfrak{A}_1, \qquad \mathfrak{N}_{1,1} = \mathfrak{E}_0$$

imply

$$(49) \qquad (\mathfrak{A}_1, \mathfrak{B}_2) = \mathfrak{E}_0, \qquad [\mathfrak{A}_2, \mathfrak{B}_1] = \mathfrak{M}.$$

From $\mathfrak{N}_{1,1} = \mathfrak{E}_0$ follows

$$\mathfrak{N}_{1,1}^{(1)} = (\mathfrak{A}_1, [\mathfrak{B}_2, (\mathfrak{A}_2, \mathfrak{B}_1)]) = \mathfrak{E}_0,$$

and hence the first condition (49) must be satisfied. Since $\mathfrak{R}_{1,1} = \mathfrak{A}_1$ we must also have $\mathfrak{N}_{1,1}^{(1)} = \mathfrak{A}_1$, and hence

$$\mathfrak{M} = [\mathfrak{N}_{1,1}^{(1)}, \mathfrak{A}_2] = [\mathfrak{A}_2, (\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{A}_1])].$$

This implies the second relation (49).

Let us now substitute the decompositions

$$\mathfrak{A}_1 = [\mathfrak{R}_{1,1}, \mathfrak{N}_{1,1}], \qquad \mathfrak{B}_1 = [\bar{\mathfrak{R}}_{1,1}, \bar{\mathfrak{N}}_{1,1}] \tag{50}$$

in the original decomposition (9). We then obtain the further direct decompositions

$$\mathfrak{M} = [\mathfrak{R}_{1,1}, \mathfrak{N}_{1,1}, \mathfrak{A}_2] = [\bar{\mathfrak{R}}_{1,1}, \bar{\mathfrak{N}}_{1,1}, \mathfrak{B}_2]. \tag{51}$$

If one takes here the component of $\mathfrak{R}_{1,1}$ in $\bar{\mathfrak{R}}_{1,1}$, one finds that it is equal to $\bar{\mathfrak{R}}_{1,1}$ by using the identities (46). In the same way $\mathfrak{R}_{1,1}$ is the $\mathfrak{R}_{1,1}$-component of $\bar{\mathfrak{R}}_{1,1}$, and $\mathfrak{R}_{1,1}$ and $\bar{\mathfrak{R}}_{1,1}$ must be their own reduction invariants with respect to each other.

This last remark in connection with Theorem 14 gives us the following relations:

$$(\mathfrak{R}_{1,1}, [\bar{\mathfrak{N}}_{1,1}, \mathfrak{B}_2]) = (\bar{\mathfrak{R}}_{1,1}, [\mathfrak{N}_{1,1}, \mathfrak{A}_2]) = \mathfrak{E}_0 \tag{52}$$

and

$$\mathfrak{M} = [\bar{\mathfrak{R}}_{1,1}, \mathfrak{N}_{1,1}, \mathfrak{A}_2] = [\mathfrak{R}_{1,1}, \bar{\mathfrak{N}}_{1,1}, \mathfrak{B}_2]. \tag{53}$$

From these results we conclude:

Theorem 15. *In the direct decompositions* (51) *for* $\mathfrak{M}$ *the two quotients* $\mathfrak{R}_{1,1}$ *and* $\bar{\mathfrak{R}}_{1,1}$ *may be interchanged to give the two new direct decompositions* (53) *for* $\mathfrak{M}$.

Another interesting fact is the following:

Theorem 16. *All null-quotients* $\mathfrak{N}_{1,2}^{(n)}$ $(n = 0, 1, \cdots)$ *are invariant with respect to taking reductions in* $\mathfrak{B}_1$,

$$R_{1,1}\{\mathfrak{N}_{12}^{(n)}\} = \mathfrak{N}_{12}^{(n)}, \tag{54}$$

and hence all of them are factors of $\mathfrak{R}_{11}$.

We shall prove (54) by induction. We observe first that

$$\mathfrak{N}_{12}^{(n)} \leqq [(\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{N}_{12}^{(n)}]), (\mathfrak{B}_2, [\mathfrak{B}_1, \mathfrak{N}_{12}^{(n)}])].$$

We take the $\mathfrak{A}_1$-component of both sides and apply (20). Since $\mathfrak{N}_{12}^{(n)}$ is a factor of $\mathfrak{A}_1$, we find

$$\mathfrak{N}_{12}^{(n)} \leqq [R_{11}\{\mathfrak{N}_{12}^{(n)}\}, (\mathfrak{N}_{12}^{(n-1)}, \mathfrak{N}_{11}^{(1)})]. \tag{55}$$

On the other hand, we have

$$(\mathfrak{B}_1, [\mathfrak{B}_2, \mathfrak{N}_{12}^{(n)}]) \leqq [\mathfrak{N}_{12}^{(n)}, (\mathfrak{B}_2, [\mathfrak{B}_1, \mathfrak{N}_{12}^{(n)}])],$$

and by again taking the $\mathfrak{A}_1$-component we find

$$R_{11}\{\mathfrak{N}_{12}^{(n)}\} \leqq [\mathfrak{N}_{12}^{(n)}, (\mathfrak{N}_{12}^{(n-1)}, \mathfrak{N}_{11}^{(1)})] = \mathfrak{N}_{12}^{(n)}.$$

Hence both terms on the r.h. side of (55) are contained in $\mathfrak{N}_{12}^{(n)}$ and the equality

sign must hold. Furthermore, since the theorem holds for $\mathfrak{N}_{12}^{(n-1)}$, the term $(\mathfrak{N}_{12}^{(n-1)}, \mathfrak{N}_{11}^{(1)})$ is contained in $R_{11}\{\mathfrak{N}_{12}^{(n-1)}\} \leqq R_{11}\{\mathfrak{N}_{12}^{(n)}\}$ and (55) is proved.

A consequence of Theorem 16 is that $\mathfrak{R}_{12}$ is a factor of $\mathfrak{R}_{11}$ and also $\mathfrak{R}_{11}$ a factor of $\mathfrak{R}_{12}$. Since we also have

$$\mathfrak{A}_1 = [\mathfrak{R}_{11}, \mathfrak{N}_{11}] = [\mathfrak{R}_{12}, \mathfrak{N}_{12}],$$

we find the further direct decompositions

$$\mathfrak{R}_{11} = [\mathfrak{N}_{12}, (\mathfrak{R}_{11}, \mathfrak{R}_{12})], \qquad \mathfrak{R}_{12} = [\mathfrak{N}_{11}, (\mathfrak{R}_{11}, \mathfrak{R}_{12})]. \tag{56}$$

THEOREM 17. *Let $\mathfrak{M}$ be a regular quotient for which there exist two direct decompositions*

$$\mathfrak{M} = [\mathfrak{A}_1, \mathfrak{A}_2] = [\mathfrak{B}_1, \mathfrak{B}_2].$$

Then for each quotient $\mathfrak{A}_1$ and $\mathfrak{A}_2$ (or $\mathfrak{B}_1$ and $\mathfrak{B}_2$) there exists a direct decomposition into three quotients

$$\mathfrak{A}_1 = [\mathfrak{N}_{11}, \mathfrak{N}_{12}, (\mathfrak{R}_{11}, \mathfrak{R}_{12})]. \tag{57}$$

7. **Proof for the main theorem.** There are still a number of interesting properties of the invariants $\mathfrak{R}$ and $\mathfrak{N}$ which we have not touched upon. Some of them are of considerable importance for the study of the properties of direct decompositions. For instance, we have seen that $\mathfrak{R}_{11}$ is the maximal direct component of $\mathfrak{A}_1$ which may be interchanged with a direct component of $\mathfrak{B}_1$ and $\mathfrak{R}_{12}$ has a similar property with respect to $\mathfrak{B}_2$. One may now ask for the maximal direct component of $\mathfrak{A}_1$ interchangeable both with a component of $\mathfrak{B}_1$ and $\mathfrak{B}_2$. This would lead us to find factors of $\mathfrak{A}_1$ which are invariant both with respect to taking reductions in $\mathfrak{B}_1$ and $\mathfrak{B}_2$. Such an investigation might be carried through along the same lines as here. The theory may also be extended to the case where $\mathfrak{M}$ is the direct union of an arbitrary number of quotients $\mathfrak{A}_i$ and $\mathfrak{B}_i$. In this case one has the possibility of taking reductions with respect to one or more components in arbitrary orders and one obtains a great number of various reduction and increment invariants with corresponding direct decompositions for the $\mathfrak{A}_i$ and $\mathfrak{B}_i$. We shall, however, not carry through the discussion of this theory.

We shall conclude these investigations by applying them to the proof of the main theorem.

THEOREM 18. *In two different direct decompositions*

$$\mathfrak{M} = [\mathfrak{A}_1, \cdots, \mathfrak{A}_r] = [\mathfrak{B}_1, \cdots, \mathfrak{B}_s] \tag{58}$$

of a quotient $\mathfrak{M}$ into direct indecomposable quotients, both sides must contain the same number of quotients directly similar in pairs.

We shall prove the theorem under the assumption that its r.h. (l.h.) factors always have regular decompositions (45). This is certainly the case when $\mathfrak{M}$ has a finite length, but it is also true for more general structures, as we shall see in the next chapter.

Let us prove first that any $\mathfrak{A}_i$ may replace a suitable $\mathfrak{B}_j$ in (58) to give a new direct decomposition of $\mathfrak{M}$. We write

$$\bar{\mathfrak{A}}_1 = [\mathfrak{A}_2, \cdots, \mathfrak{A}_r], \qquad \bar{\mathfrak{B}}_1 = [\mathfrak{B}_2, \cdots, \mathfrak{B}_s],$$

and we shall prove that $\mathfrak{A}_1$ may replace some $\mathfrak{B}_j$. Since $\mathfrak{A}_1$ is directly indecomposable, its reduction and increment invariants with respect to any quotient are either $\mathfrak{A}_1$ or $\mathfrak{E}_0$. If the reduction invariant of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$ in (58) is equal to $\mathfrak{A}_1$, then $\mathfrak{A}_1$ and $\mathfrak{B}_1$ are interchangeable according to Theorem 15. Hence we may suppose that the reduction invariant of $\mathfrak{A}_1$ with respect to $\mathfrak{B}_1$ is $\mathfrak{E}_0$. According to (56) we then have that the null-quotient of $\mathfrak{A}_1$ with respect to $\bar{\mathfrak{B}}_1$ must be $\mathfrak{E}_0$, and this is easily seen to imply $(\mathfrak{A}_1, \mathfrak{B}_1) = \mathfrak{E}_0$.

From Theorem 13 it follows that in our case the reduction invariant of $\mathfrak{A}_1$ with respect to $\bar{\mathfrak{B}}_1$ is $\mathfrak{A}_1$, while the corresponding increment invariant is $\mathfrak{E}_0$. The corresponding quotients for $\bar{\mathfrak{B}}_1$ with respect to $\mathfrak{A}_1$ are then found by Theorem 12,

$$\mathfrak{R} = (\bar{\mathfrak{B}}_1, [\mathfrak{B}_1, \mathfrak{A}_1]), \qquad \mathfrak{R} = (\bar{\mathfrak{B}}_1, \bar{\mathfrak{A}}_1),$$

where $\mathfrak{R}$ is directly indecomposable because it is similar to $\mathfrak{A}_1$. This gives us the direct decomposition

$$\bar{\mathfrak{B}}_1 = [(\bar{\mathfrak{B}}_1, [\mathfrak{B}_1, \mathfrak{A}_1]), (\bar{\mathfrak{B}}_1, \bar{\mathfrak{A}}_1)] = [\mathfrak{B}_2, \cdots \mathfrak{B}_s]. \tag{59}$$

If we now suppose we have proved that any quotient $\mathfrak{A}_i$ in a decomposition (58) may replace some $\mathfrak{B}_j$ when the r.h. decomposition contains less than s terms, we may apply this result to (59). When $\mathfrak{R}$ may replace $\mathfrak{B}_2$, we find

$$\bar{\mathfrak{B}}_1 = [(\bar{\mathfrak{B}}_1, [\mathfrak{B}_1, \mathfrak{A}_1]), \mathfrak{B}_3, \cdots, \mathfrak{B}_s],$$

and by taking the union with $\mathfrak{B}_1$, one obtains the new direct decomposition

$$\mathfrak{M} = [\mathfrak{B}_1, \mathfrak{A}_1, \mathfrak{B}_3, \cdots, \mathfrak{B}_s],$$

showing that $\mathfrak{A}_1$ may replace $\mathfrak{B}_2$.

The proof of Theorem 18 is now simple. We suppose that $\mathfrak{A}_1$ may replace $\mathfrak{B}_1$,

$$\mathfrak{M} = [\mathfrak{B}_1, \cdots, \mathfrak{B}_s] = [\mathfrak{A}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_s].$$

This shows that $\mathfrak{A}_1$ and $\mathfrak{B}_1$ are directly similar. The quotient

$$\mathfrak{M} \times \mathfrak{A}_1^{-1} = [\mathfrak{A}_1, \mathfrak{A}_2, \mathfrak{A}_1^{-1}, \cdots] = [\mathfrak{A}_1, \mathfrak{B}_2, \mathfrak{A}_1^{-1}, \cdots] \tag{60}$$

is then a l.h. factor of $\mathfrak{M}$, but it is also similar to the r.h. factor $\bar{\mathfrak{A}}_1$. Using induction, we may assume that the theorem is true for the quotient (60); hence we have $r = s$ and the direct similarity of the other quotients $\mathfrak{A}_i$ and $\mathfrak{B}_i$ is easily obtained.

Chapter 2. Conditions for the main theorem

1. **Existence of the invariants.** The theory which we have derived in Chapter 1 depends entirely upon the existence of the decomposition (45) of a

quotient $\mathfrak{A}_1$ into its reduction and increment invariants. Even when no finiteness condition of any kind is imposed upon the quotient $\mathfrak{M}$ it is possible to define reduction and increment invariants under very general conditions, but it seems difficult in this case to draw any conclusions about the existence of direct decompositions of the form (41). Hence we shall have to suppose in the following that either the ascending or the descending chain condition is satisfied in $\mathfrak{M}$.

We shall use the same notation as in Chapter 1, considering a direct decomposition

$$\mathfrak{M} = [\mathfrak{A}_1, \mathfrak{A}_2] = [\mathfrak{B}_1, \mathfrak{B}_2]. \tag{1}$$

We may drop all subscripts of the occurring quotients, since we shall only consider the reductions of $\mathfrak{A}_1$ in $\mathfrak{B}_1$. In this simplified notation let us recall that the reduction invariant $\mathfrak{R}$ and the increment invariant $\mathfrak{N}$ were defined by

$$\mathfrak{R} = \mathfrak{R}^{(n)} = \mathfrak{R}^{(n+1)} = \cdots \tag{2}$$

and

$$\mathfrak{N} = \mathfrak{N}^{(m)} = \mathfrak{N}^{(m+1)} = \cdots . \tag{3}$$

Let us make the preliminary observation that when $\mathfrak{R}$ and $\mathfrak{N}$ exist, we have $n = m$ for the smallest indices n and m for which (1) and (2) hold. When $\mathfrak{R}$ exists, we have according to Theorem 10, Chapter 1,

$$\mathfrak{A}_1 = [\mathfrak{R}, \mathfrak{N}^{(n)}], \tag{4}$$

and by taking the n-th reduction of both sides we obtain $\mathfrak{N}^{(m)} = \mathfrak{N}$ and hence $m \geqq n$. Similarly it follows from

$$\mathfrak{E}_0 = (\mathfrak{N}, \mathfrak{R}^{(m)}) \tag{5}$$

that $n \geqq m$.

Let us now suppose that the ascending chain condition holds in $\mathfrak{M}$ (or only in $\mathfrak{A}_1$). The increment invariant $\mathfrak{N}$ must then exist and the relation (5) holds. Our problem is to determine the conditions for the existence of $\mathfrak{R}$. According to the preceding remark, the necessary and sufficient condition for this is

$$\mathfrak{R}^{(m)} = \mathfrak{R}^{(m+1)} \tag{6}$$

or, as one easily sees,

$$\mathfrak{A}_1 = [\mathfrak{R}^{(m)}, \mathfrak{N}]. \tag{7}$$

From Theorem 6, Chapter 1 it follows that we always have the similarity relation

$$\mathfrak{R}^{(m+1)} \cong \mathfrak{R}^{(m)}/(\mathfrak{R}^{(m)}, \mathfrak{N}^{(1)}).$$

Since $\mathfrak{R}^{(m)}$ is relatively prime to $\mathfrak{N}$ according to (5), it is also relatively prime to $\mathfrak{N}^{(1)}$, and hence we obtain

$$\mathfrak{R}^{(m+1)} \cong \mathfrak{R}^{(m)}. \tag{8}$$

This proves that if $\mathfrak{A}_1$ has no factors $\mathfrak{D}$ and $\mathfrak{D}' < \mathfrak{D}$ such that $\mathfrak{D} \cong \mathfrak{D}'$, the regular decomposition

$$\mathfrak{A}_1 = [\mathfrak{N}, \mathfrak{R}], \qquad (\mathfrak{N}, \mathfrak{R}) = \mathfrak{E}_0 \tag{9}$$

must exist.

Similarly, if the descending chain condition holds in $\mathfrak{A}_1$, the reduction invariant $\mathfrak{R}$ must exist and we find corresponding to (8)

$$\mathfrak{A}_1/\mathfrak{R}^{(m+1)} \cong \mathfrak{A}_1/\mathfrak{R}^{(m)}. \tag{10}$$

Hence in this case the regular decomposition (9) must exist provided $\mathfrak{A}_1$ has no left-hand factor $\mathfrak{D}$ with a proper l.h. factor $\mathfrak{D}'$ such that $\mathfrak{D} \cong \mathfrak{D}'$.

Theorem 1. *Let $\mathfrak{M}$ be a quotient satisfying the ascending (descending) chain condition. Then the regular decomposition* (9) *will exist and hence the main theorem about direct decompositions will hold provided $\mathfrak{M}$ has no right-hand (l.h.) factor $\mathfrak{D}$ containing a proper r.h. (l.h.) factor $\mathfrak{D}''$ such that $\mathfrak{D}$ and $\mathfrak{D}'$ are similar.*

One may, however, improve considerably upon Theorem 1 by observing that the similarity relations (8) and (10) are of a very special nature. Let us suppose again that the ascending chain condition holds in $\mathfrak{A}_1$ and let us denote by $\mathfrak{B}^{(m)}$ the component of $\mathfrak{R}^{(m)}$ in $\mathfrak{B}_1$. We find then

$$[\mathfrak{R}^{(m)}, \mathfrak{B}_2] = [\mathfrak{B}^{(m)}, \mathfrak{B}_2], \tag{11}$$

where

$$(\mathfrak{R}^{(m)}, \mathfrak{B}_2) = (\mathfrak{B}^{(m)}, \mathfrak{B}_2) = \mathfrak{E}_0. \tag{12}$$

The last relation (12) is obvious, because $\mathfrak{B}^{(m)}$ is a factor of $\mathfrak{B}_1$ and the first follows from the fact that $\mathfrak{R}^{(m)}$ is relatively prime to $\mathfrak{R}^{(1)}$ according to (5).

The component of $\mathfrak{B}^{(m)}$ in $\mathfrak{A}_1$ is $\mathfrak{R}^{(m+1)}$ and we find as before that

$$[\mathfrak{R}^{(m+1)}, \mathfrak{A}_2] = [\mathfrak{B}^{(m)}, \mathfrak{A}_2], \tag{13}$$

where we also have

$$(\mathfrak{R}^{(m+1)}, \mathfrak{A}_2) = (\mathfrak{B}^{(m)}, \mathfrak{A}_2) = \mathfrak{E}_0. \tag{14}$$

To prove the last relation it is only necessary to observe that $(\mathfrak{A}_2, \mathfrak{B}^{(m)})$ is found to be the $\mathfrak{B}_1$-component of

$$(\mathfrak{R}^{(m)}, \mathfrak{R}^{(1)}) = \mathfrak{E}_0.$$

We have formerly defined two quotients $\mathfrak{A}$ and $\mathfrak{B}$ to be directly similar when there exists a third quotient $\mathfrak{C}$ relatively prime to both $\mathfrak{A}$ and $\mathfrak{B}$ such that

$$[\mathfrak{A}, \mathfrak{C}] = [\mathfrak{B}, \mathfrak{C}].$$

The notion of direct similarity in a Dedekind structure is usually not transitive. This leads us to introduce another special type of similarity: two quotients $\mathfrak{A}$ and $\mathfrak{B}$ are said to be *semi-directly similar* when there exists a $\mathfrak{C}$ to which they are both directly similar.

The relations (11), (12) and (13), (14) show that $\mathfrak{R}^{(m)}$ and $\mathfrak{R}^{(m+1)}$ are semi-

directly similar. One may define left-hand semi-direct similarity in a corresponding manner, and one finds naturally that when $\mathfrak{N}$ exists, the two quotients $\mathfrak{A}_1/\mathfrak{N}^{(n)}$ and $\mathfrak{A}_1/\mathfrak{N}^{(n+1)}$ are l.h. semi-directly similar.

THEOREM 2. *The main theorem holds in $\mathfrak{M}$ when the ascending (descending) chain condition is satisfied and $\mathfrak{M}$ contains no r.h. (l.h.) factor $\mathfrak{D}$ with a proper r.h. (l.h.) factor $\mathfrak{D}'$ such that $\mathfrak{D}$ and $\mathfrak{D}'$ are semi-directly similar.*

2. **Axiomatic conditions.** The preceding results naturally lead us to the consideration of structures having the following special property:

I. (r.h.). *When $\mathfrak{A} \geqq \mathfrak{A}'$ and $\mathfrak{A}$ and $\mathfrak{A}'$ are r.h. semi-directly similar, then $\mathfrak{A} = \mathfrak{A}'$.*

We may say that a structure Σ in which this condition is satisfied is *r.h. semi-directly regular.* One may also express the condition for r.h. semi-direct regularity in the following manner:

I$'$(r.h.). *Let $A \geqq A'$ be two elements in the structure Σ. If the relations*

$$(15)\qquad \begin{gathered} [A, B] = [C, B], \qquad\qquad [C, D] = [A', D] \\ (A, B) = (C, B) = (C, D) = (A', D) \end{gathered}$$

hold, we can conclude $A = A'$.

The dualistically corresponding condition for l.h. semi-direct regularity is obviously

I$'$(l.h.). *If we have $A \geqq A'$ and the relations*

$$(16)\qquad \begin{gathered} [A, B] = [C, B] = [C, D] = [A', D] \\ (A, B) = (C, B), \qquad\qquad (C, D) = (A', D), \end{gathered}$$

we can conclude $A = A'$.

In the last formulations the condition for semi-direct regularity reminds one strikingly of the following formulations of the two principal axioms.[5]

DISTRIBUTIVE AXIOM. *If*

$$(17)\qquad [A, B] = [A', B], \qquad (A, B) = (A', B),$$

then $A = A'$.

DEDEKIND AXIOM. *If $A \geqq A'$ and the relations* (17) *hold, we can conclude $A = A'$.*

It is obvious that the distributive axiom implies semi-direct regularity, since in distributive structures direct similarity implies equality. More interesting is the fact that either r.h. or l.h. semi-direct regularity implies the Dedekind axiom. To prove this we need only make $B = D$ and $C = A$ in (15) or (16).

THEOREM 3. *The distributive axiom implies semi-direct regularity and semi-direct regularity implies the Dedekind axiom.*

I have formerly proved the main theorem on direct decompositions in Dede-

[5] See Ore I, Chap. 1.

kind structures, where the descending chain condition holds and where in addition the following axiom is satisfied:[6]

II. (l.h.). *Let* $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ *and* $\mathfrak{D}$ *be four quotients with the same denominator such that*

$$[\mathfrak{A}, \mathfrak{B}] = [\mathfrak{C}, \mathfrak{D}] = [\mathfrak{C}, \mathfrak{B}] = [\mathfrak{A}, \mathfrak{D}].$$

If then $(\mathfrak{A}, \mathfrak{B}) = (\mathfrak{C}, \mathfrak{D}) = \mathfrak{E}_0$, *we can conclude* $(\mathfrak{B}, \mathfrak{C}) = (\mathfrak{A}, \mathfrak{D}) = \mathfrak{E}_0$.

We shall say that a structure Σ in which II (l.h.) is satisfied is (l.h.) *weakly regular*. The condition for weak regularity may also be stated:

II′ (l.h.). *If the relations*

$$M = [A, B] = [C, D] = [C, B],$$
$$T = (A, B) = (C, D), (C, B) \geqq T, (A, D) \geqq T$$

are satisfied, we can conclude $T = (A, B) = (C, D) = (C, B) = (A, D)$.

Correspondingly we have

II′ (r.h.). *If*

$$T = (A, B) = (C, D) = (C, B) = (A, D),$$
$$M = [A, B] = [C, D], [C, B] \leqq M, [A, D] \leqq M,$$

we can conclude $M = [A, B] = [C, D] = [C, B] = [A, D]$.

Theorem 4. *Right-hand (l.h.) semi-direct regularity implies r.h. (l.h.) weak regularity.*

Let us suppose that the conditions of II′ (r.h.) are satisfied. The relations

$$[A, D] = [(C, [A, D]), D], [(C, [A, D]), B] = [(A, [B, (C, [D, A])]), B]$$

show that the quotients A/T and A'/T, where $A' = (A, [B, (C, [D, A])])$ are semi-directly similar. Hence if Σ is semi-directly regular we have $A = A'$, so that

$$M = [B, (C, [D, A])]$$

and consequently $M = [B, C]$. The relation $M = [A, D]$ is proved in a similar manner.

The decomposition theory of Chapter 1 is valid when the descending (ascending) chain condition holds and Σ is l.h. (r.h.) semi-directly regular. On the other hand, the main theorem on direct decompositions has been proved for Dedekind structures where the descending (ascending) chain condition holds and which are l.h. (r.h.) weakly regular. I have not been able to carry through the general decomposition theory under these weaker conditions and it seems possible that the existence of the decompositions of Chapter 1 requires a somewhat stronger axiom than the main theorem. It seems an interesting problem to be considered.

Yale University

[6] Ore II, Chap. 2.

SEMI-CONTINUITY OF INTEGRALS IN THE CALCULUS OF VARIATIONS

By E. J. McShane

Introduction. In various studies of existence theorems in the calculus of variations much use has been made of the property of lower semi-continuity of the integrals involved. For each separate type of problem there has been a separate proof of semi-continuity. The principal object of this paper is to prove one theorem on semi-continuity of integrals which has generality enough to cover as special cases the simple problem in parametric form and in ordinary form, the Lagrange problem in parametric form and in ordinary form, and the parametric problem associated with a problem in ordinary form.[1] As a by-product we are able to state existence theorems for certain problems not covered by the existence theorems in the literature.

The purpose of §2 is merely to extend to our analytical situation the every-day theorems in invariance under change of parameter. In §3 the principal semi-continuity theorem is proved. The notation is that of the parametric problem, but the hypotheses are so weak as to permit us in §4 to restate it in ordinary form. In §5 it is specialized to cover parametric problems and Lagrange problems in parametric form. In §6 it is specialized to cover Lagrange problems in ordinary form. The next section gives three examples to indicate that the hypotheses in §6 do not admit of much weakening. One of these examples (Example III) has an interest quite apart from semi-continuity theory, for in it we exhibit a Lagrange problem in ordinary form for which $y \equiv 0$ is an extremal imbedded in a field of extremals, furnishing a weak relative minimum for the integral in the class of admissible curves, satisfying the Legendre and Weierstrass conditions along $y \equiv 0$ (but not, of course, in strengthened form) and yet $y \equiv 0$ does not afford a strong relative minimum for the integral. In §8 we deduce from the general theorem a theorem on the semi-continuity of the parametric integral $\int f(x, y, x', y')dt$ associated with a problem $\int g(x, y, y')dx$ in ordinary form. The specializations in §§5, 7, and 8 yield Theorems 5.1, 6.1, 6.2, 8.1, 8.2, which, to the best of my knowledge, are stronger than any in the literature.

If to the hypotheses of the semi-continuity theorem we add the hypothesis that the integrand is positive, we can easily obtain an existence theorem. In §9 we apply this existence theorem to three special cases. The first is that of finding the path of a beam of light through a space in which pieces of glass are suspended. The second is the Zermelo navigation problem.[2] Here the inte-

Received June 24, 1936.

[1] E. J. McShane, *Existence theorems for ordinary problems*, etc., Annali della R. Sc. Norm. Sup. di Pisa, Ser. II, vol. III (1934), pp. 183–211, pp. 287–315.

[2] Carathéodory, *Variationsrechnung*, p. 234.

grand $f(x, x')$ is not defined for all x', and it is useful that our semi-continuity theorem has hypotheses so weak as to permit us to set $f(x, x') = +\infty$ where it is not already defined. The third problem is the special case of the Lagrange problem in parametric form in which the side equations are linear.

All of our semi-continuity theorems state merely that, under appropriate conditions, the integral involved is lower semi-continuous on a class of curves of uniformly bounded lengths. This restriction is really not so important, since in establishing our existence theorem we need not only a semi-continuity theorem but also a theorem establishing the convergence of a minimizing sequence of curves, and in order to establish this convergence it seems vital to have a uniform bound on the lengths of the curves in question. However, from the semi-continuity theorems here obtained it would be very easy to obtain conditions guaranteeing semi-continuity on the class of all admissible curves, without a uniform bound on the lengths.[3]

§1. **Notation and definitions.** The letters x, y, r, η will be used to denote vectors; y will stand for $(y^1, y^2, \cdots, y^q)$, η for $(\eta^1, \eta^2, \cdots, \eta^q)$, while x, r will stand for $(x^0, \cdots, x^q)$, $(r^0, \cdots, r^q)$, respectively. If a function $f(x, r)$ has a partial derivative with respect to r^i, that derivative will be denoted by $f_{(i)}(x, r)$. Likewise, if $g(u, y, \eta)$ has a partial derivative with respect to η^i, that derivative will be denoted by $g_{(i)}(u, y, \eta)$. The lengths of the vectors x, y, r, η will be denoted by $|x|$, $|y|$, $|r|$, $|\eta|$ respectively. We use a modification of the tensor summation; if a *Greek-letter* affix is repeated, the expression is to be summed over all values of that affix. Thus

$$r_n^\alpha f_{(\alpha)}(x_n, r_n) = r_n^0 f_{(0)}(x_n, r_n) + \cdots + r_n^q f_{(q)}(x_n, r_n),$$

summed on α but not on n.

Functions will be permitted to assume the value of $+\infty$, but not $-\infty$. For the symbol ∞ we use the rules of calculation $\infty + \infty = \infty + a = a + \infty = \infty$ for all finite numbers a; $a\infty = \infty$ if $a > 0$, $0\,\infty = 0$. These rules cover all cases which will occur. The notion of lower semi-continuity extends, of course, to such functions (if we take $\infty > a$ for all finite a); a function $f(x)$ is lower semi-continuous (hereafter abbreviated to l.s.c.) on a set E if for every x_0 in E and every number $h < f(x_0)$ there is a neighborhood U of x_0 such that $f(x) > h$, for $x \in EU$.

The integrals used will be Lebesgue integrals, with one minor modification; if a function $\phi(x)$ is measurable but not summable over a set E, and there is a

[3] In terms of the notation about to be introduced, the additional requirement is that for each (x_0, r_0) in R there shall be a linear function $b_\alpha r^\alpha$ and an $\epsilon > 0$, such that $f(x, r) + b_\alpha r^\alpha \geqq \epsilon\,|r|$ for all x near x_0. This condition can be deduced from various other conditions in the special problems considered. For instance, in the parametric problem $\int f(x, x')dt = \min$, if the integrand has partial derivatives $f_{(i)}(x, r)$ continuous for $|r| \neq 0$, it is enough to add the assumption that $\mathcal{E}(x, r, \bar{r})$ does not vanish identically for any x.

summable function $g(x)$ such that $\phi(x) \geqq g(x)$, we shall define

$$\int_E \phi(x)dx = \infty.$$

The letters a.c. will be used in place of the words "absolutely continuous". Given a function, for example $x(t)$, the symbol $\dot{x}(t)$ shall denote the derivative $x'(t)$, where $x'(t)$ is defined and finite, and shall have the value 0 elsewhere.

§2. Since our goal is a semi-continuity theorem which, among other things, covers the Lagrange problem, it is appropriate for us to suppose that at each point of our space there is a restriction on the set of directions which may be taken by the curves which we wish to admit. Correspondingly, it is desirable that the conditions imposed on our integrand shall refer only to these allowable directions. Guided by these considerations and our needs in the following pages, we set down the following conditions on our integrand $f(x, r)$:

(2.1a) $f(x, r)$ is defined (finite or $+\infty$) and l.s.c. on a set R in (x, r)-space;
(2.1b) if $(x, r) \in R$, then $(x, tr) \in R$, and $f(x, tr) = tf(x, r)$ for all $t \geqq 0$;
(2.1c) R is dense in itself, and the set of x such that $(x, 0) \in R$ is closed;
(2.1d) if $(x_0, r_0) \in R$ and $u < f(x_0, r_0)$, then there exists a linear function $a_\alpha r^\alpha$ with the properties (i) $a_\alpha r_0^\alpha > u$ and (ii) for every $\epsilon > 0$ there is a neighborhood U of x_0 such that $f(x, r) \geqq a_\alpha r^\alpha - \epsilon \mid r \mid$ whenever $x \in U$ and $(x, r) \in R$.

We can, however, state another set of conditions, slightly more restrictive, but satisfied by nearly all the integrands we shall discuss:

(2.2a) $f(x, r)$ is defined (finite or $+\infty$) and l.s.c. on the closure $\bar{R}$ of a set R in (x, r)-space;
(2.2b) if $(x, r) \in R$, then $(x, tr) \in R$ and $f(x, tr) = tf(x, r)$ for all $t \geqq 0$;
(2.2c) R is dense in itself, and the set of x such that $(x, 0) \in R$ is closed;
(2.2d) if $(x_0, r_0) \in R$ and $u < f(x_0, r_0)$, there exists a linear function $a_\alpha r^\alpha$ such that (i) $u < a_\alpha r_0^\alpha$ and (ii) $a_\alpha r^\alpha \leqq f(x_0, r)$ for all $(x_0, r) \in \bar{R}$.

Here it is obvious that (2.1, a, b, c) follow from (2.2). To obtain (2.1d), we note that if the linear function $a_\alpha r^\alpha$ of (2.2d) does not satisfy (2.1d), then for some $\epsilon > 0$ there is a sequence (x_n, r_n) of elements of R with $x_n \to x_0$ such that $f(x_n, r_n) < a_\alpha r_n^\alpha - \epsilon \mid r_n \mid$. By (2.2b) we may suppose $\mid r_n \mid = 1$. From the (x_n, r_n) we select a subsequence (x_p, r_p) such that r_p tends to a limit $\bar{r}$. Then $(x_0, \bar{r}) \in \bar{R}$, and by (2.2a) we find $f(x_0, \bar{r}) \leqq \liminf f(x_p, r_p) \leqq a_\alpha \bar{r}^\alpha - \epsilon \mid \bar{r} \mid < a_\alpha \bar{r}^\alpha$, contradicting (2.2d).

Of the conditions (2.2), (a) and (b) are obvious weakenings of standard hypotheses. Also, the requirement that R be dense in itself is almost trivial, for an isolated point of R could not lie on any admissible[4] curve except a degenerate curve consisting of one point, and so may be disregarded. With (2.2d) it is

[4] This term will be defined in the next paragraph.

different; this is our "regularity condition". If, in particular, it happens that R is closed, that $f(x, r)$ has partial derivatives $f_{(i)}(x, r)$ with respect to the r^i and that

$$\mathcal{E}(x, r, \bar{r}) \equiv f(x, \bar{r}) - \bar{r}^\alpha f_{(\alpha)}(x, r) \geqq 0$$

whenever (x, r) and $(x, \bar{r})$ are in R and $|r| \neq 0$, then (2.2d) is satisfied if we take $a_i = f_{(i)}(x_0, r_0)$. Again, if R consists of all sets (x, r) with x in a set A and r arbitrary, then (2.2d) is satisfied if and only if $f(x, r)$ is a convex function of r for each fixed x.

A representation $x = x(t)$, $a \leqq t \leqq b$, will be called *admissible* if the functions $x(t)$ are absolutely continuous and $(x(t), \dot{x}(t)) \in R$ for almost all t. We shall now prove that (under conditions 2.1) if $x = x(t)$, $a \leqq t \leqq b$, is any a.c. representation of a curve C and $x = \xi(s)$, $0 \leqq s \leqq L$, is the representation of C with arc-length as parameter, then $x = x(t)$ is admissible if and only if $x = \xi(s)$ is admissible. Since

$$\xi^i(s) = \xi^i(0) + \int_0^s \dot{\xi}^i(s)ds,$$

and $s(t)$ is a.c., we find

$$x^i(t) = \xi^i(s(t)) = \xi^i(0) + \int_a^t \dot{\xi}^i(s(t))\dot{s}(t)dt,$$

so that

$$\dot{x}^i(t) = \dot{\xi}^i(s)\,\dot{s}(t) \tag{2.3}$$

for almost all t. Suppose now that $x = x(t)$ is admissible. Let T_0 be the set (of measure 0) on which (2.3) fails or $(x(t), \dot{x}(t))$ is not in R, and let T_1 be the set on which $\dot{s}(t) = 0$. The measure of the set $s(t)$, $t \in (T_0 + T_1)$ is[5]

$$\int_{T_0+T_1} \dot{s}(t)dt = 0.$$

For all other s we have $\dot{\xi}(s) = \dot{x}(t) \div \dot{s}(t)$ with $(x, \dot{x}) \in R$ and $\dot{s} > 0$, so by (2.1b) we see that $(\xi(s), \dot{\xi}(s)) \in R$, and $x = \xi(s)$ is admissible. On the other hand, suppose that $x = \xi(s)$ is admissible. Then $(\xi(s), \dot{\xi}(s)) \in R$ for all values of s except those in a set S_0 of measure 0. Let T_0 be the set such that $s(t) \in S_0$ for $t \in T_0$. For almost all t not in T_0 we have $\dot{x}(t) = \dot{\xi}(s(t))\dot{s}(t)$, so $(x(t), \dot{x}(t)) \in R$ by (2.1b). For almost all t in T_0 we have[6] $\dot{s}(t) = 0$ and $\dot{x}(t) = \dot{\xi}(s(t))\cdot 0 = 0$. Since $(x(\tau), \dot{x}(\tau))$ is in R for almost all τ, so is $(x(\tau), 0)$. We can choose a sequence of τ tending to t; then $x(\tau) \to x(t)$, and by (2.1c) the set $(x(t), 0) = (x(t), \dot{x}(t))$ is in R. Hence for almost all t the set $(x(t), \dot{x}(t))$ is in R and $x = x(t)$ is an admissible representation.

[5] Hobson, *Theory of Functions of a Real Variable*, vol. I, pp. 606 and 342.

[6] S_0 is contained in a G_δ-set S of measure 0. The set T on which $s(t) \in S$ is also a G_δ, hence measurable. By footnote 5, $0 = mS = \int_T \dot{s}(t)dt$, so $\dot{s}(t) = 0$ almost everywhere in T and à fortiori almost everywhere in T_0.

It follows at once that if one representation of a curve C is admissible, so also are all other a.c. representations. Hence in this case we are justified in saying that C is an admissible curve. We shall understand that all representations of curves hereafter mentioned are a.c., so that if C is an admissible curve $x = x(t)$, the representation $x = x(t)$ is admissible.

We can now prove

THEOREM 2.1. *If $x = x(t)$, $a \leqq t \leqq b$, and $x = \bar{x}(\tau)$, $\alpha \leqq \tau \leqq \beta$, are two representations of the same admissible curve C, then the integrals*

$$\int_a^b f(x, \dot{x})dt \quad \textit{and} \quad \int_\alpha^\beta f(\bar{x}, \dot{\bar{x}})d\tau$$

are both defined (finite or $+\infty$) and are equal.

Proof. As in the preceding proof, the set $(x(t), 0)$ is admissible for all t. Hence to each x_0 on C there is, by 2.1d, a linear function $a_\alpha r^\alpha$ and a neighborhood U such that $f(x, r) \geqq a_\alpha r^\alpha - |r|$ if $(x, r) \in R$ and $x \in U$. A finite number of these neighborhoods U cover the set of points $x(t)$, $a \leqq t \leqq b$. Denoting by $N - 1$ the greatest of the corresponding numbers (vector-lengths) $|a_i|$, we have $f(x, r) \geqq -N|r|$ for $(x, r) \in R$ and x in a neighborhood of the point-set $x(t)$, $a \leqq t \leqq b$. For $n > N$ we define $f_n(x, r) = \min (f(x, r), n|r|)$, $(x, r) \in R$. Then $|f_n(x, r)| \leqq n|r|$ if x is on C and $(x, r) \in R$. The function $f_n(x, r)$ is l.s.c. on R, being the minimum of two l.s.c. functions; hence $f_n(x(t), \dot{x}(t))$ is measurable.[7] The same is true of $f_n(\xi(s), \dot{\xi}(s))$, where $x = \xi(s)$ is the representation of C with arc-length as parameter. Since the functions $n|\dot{x}(t)|$ and $n|\dot{\xi}(s)|$ are summable, so are $f_n(x, \dot{x})$ and $f_n(\xi, \dot{\xi})$, and by (2.1b) and (2.3)

$$\begin{aligned}\int_0^L f_n(\xi(s), \dot{\xi}(s))ds &= \int_a^b f_n(\xi(s(t)), \dot{\xi}(s(t)))\,\dot{s}(t)dt \\ &= \int_a^b f_n(x(t), \dot{x}(t))dt.\end{aligned}$$

Now let $n \to \infty$. Then for all s and all t the functions $f_n(x(t), \dot{x}(t))$ and $f_n(\xi(s), \dot{\xi}(s))$ increase monotonically and tend respectively to $f(x(t), \dot{x}(t))$ and $f(\xi(s), \dot{\xi}(s))$. Hence the two integrals

$$\int_0^L f(\xi(s), \dot{\xi}(s))ds \quad \text{and} \quad \int_a^b f(x(t), \dot{x}(t))dt$$

both exist (finite or infinite) and are equal. By repeating the argument with $\bar{x}$ in place of x, the theorem is established.

We are now justified in denoting the common value of the integrals in Theorem 2.1 by the symbol $\mathcal{F}(C)$.

[7] Carathéodory, *Vorlesungen über Reelle Funktionen*, p. 377. The theorem does not apply at once, but $f_n(x, r) + N|r|$ is non-negative and l.s.c. for $(x, r) \in R$ and x in a neighborhood of the point-set composing C, so $f_n(x, r) + N|r|$ can be extended to be l.s.c. on all of space. This implies the measurability of $f_n(x, \dot{x}) + N|\dot{x}|$, hence of $f_n(x, \dot{x})$.

§3. These preliminaries being disposed of, we proceed to the proof of our principal theorem on semi-continuity.

THEOREM 3.1. *If hypotheses* (2.1) *are satisfied and* M *is any positive number, then* $\mathcal{F}(C)$ *is l.s.c. on the class of all admissible curves of length* $\leqq M$.

We prove this theorem in several steps.

LEMMA 3.2. *To establish Theorem* 3.1, *it is sufficient to show that for every* $M > 0$ *the integral* $\int f(x, \dot{x})dt$ *is l.s.c. on the class of all admissible functions*[8] $x = x(t)$, $0 \leqq t \leqq 1$, *such that* $|x(t)| \leqq M$ *and* $|\dot{x}(t)| \leqq M$.

Proof. Suppose Theorem 3.1 false. We can then find a number M_1 and a sequence $\{C_n\}$ of admissible curves tending to an admissible limit curve C_0, having lengths $\leqq M_1$, and satisfying the inequality $\liminf \mathcal{F}(C_n) < \mathcal{F}(C_0)$. For each curve C_n we choose as parameter $t = s/\mathfrak{L}(C_n)$, where s is the arc-length and $\mathfrak{L}(n)$ is the total length of C_n; then C_n is represented by equations $x = x_n(t)$, $0 \leqq t \leqq 1$, where $x_n(t)$ satisfies a Lipschitz condition of constant M_1 and $|\dot{x}_n(t)| \leqq M_1$. From the C_n we first choose a subsequence $\{C_\alpha\}$ such that $\lim \mathcal{F}(C_\alpha)$ exists and is equal to $\liminf \mathcal{F}(C_n)$, and then from the $\{C_\alpha\}$ we choose a subsequence $\{C_\beta\}$ such that $x_\beta(t)$ converges uniformly to a limit function $x_0(t)$; this last is possible by Ascoli's theorem. Since $x_\beta(t) \rightrightarrows x_0(t)$, the curves C_β tend to the curve represented by $x = x_0(t)$. But $\lim C_\beta = C_0$; therefore $x = x_0(t)$ is a representation of C_0. Clearly $x_0(t)$ also satisfies a Lipschitz condition of constant M_1. Since $x_\beta(t) \rightrightarrows x_0(t)$, the numbers $|x_\beta(t)|$ are bounded, say $\leqq M_2$. Setting $M = \max(M_1, M_2)$, we have $|x_\beta(t)| \leqq M$, $|\dot{x}_\beta(t)| \leqq M$, and by Theorem 2.1

$$\liminf \int_0^1 f(x_\beta, \dot{x}_\beta)\, dt < \int_0^1 f(x_0, \dot{x}_0)\, dt.$$

Hence if Theorem 3.1 is false, there is an M such that $\int f(x, \dot{x})dt$ is not l.s.c. on the class of admissible functions $x(t)$ with $|x| \leqq M$ and $|\dot{x}| \leqq M$, and our lemma is established.

The use of this lemma is that it enables us to consider the representations as fixed; and having no further need for invariance under change of parameters, we may use auxiliary functions which do not satisfy (2.1b).

LEMMA 3.3. *If hypotheses* (2.1) *are satisfied, there exists a function* $F(x, r)$, *defined and l.s.c. for all* x *and all* r, *satisfying the equation* $F(x, r) = f(x, r)$ *for* $(x, r) \in R$ *and such that if* $(x_0, r_0) \in R$ *and* $u < F(x_0, r_0)$, *there exists a linear function* $a_\alpha r^\alpha$ *for which* $a_\alpha r_0^\alpha > u$ *and* $a_\alpha r^\alpha \leqq F(x_0, r)$ *for all* r.

Proof. Let us first set $g(x, r) = f(x, r)$ for $(x, r) \in R$, and $g(x, r) = \infty$ elsewhere. Now we define $F(x, r)$ to be the lower limit function of g; that is, the smaller of $g(x, r)$ and $\liminf g(\bar{x}, \bar{r})$ as $(\bar{x}, \bar{r}) \to (x, r)$. Then F is l.s.c.[9] for all (x, r). If $(x_0, r_0) \in R$, then for every $h < f(x_0, r_0)$ there exists a neighborhood U of (x_0, r_0) such that $f(\bar{x}, \bar{r}) > h$ for $(\bar{x}, \bar{r}) \in RU$. Therefore, $g(\bar{x}, \bar{r}) > h$ for $(\bar{x}, \bar{r}) \in U$,

[8] The distance between two functions $x(t)$ and $x_1(t)$, $0 \leqq t \leqq 1$, is here understood to be $\max |x_1(t) - x(t)|$. This is a special case of a definition which will be given in §4.

[9] Carathéodory, *Vorlesungen über Reelle Funktionen*, p. 137.

and lim inf $g(\bar{x}, \bar{r}) \geqq f(x_0, r_0) = g(x_0, r_0)$. By the definition of F we then have $F(x_0, r_0) = f(x_0, r_0)$. Finally, let $a_\alpha r^\alpha$ be the linear function of (2.1d). From (2.1d) we know that for every $\epsilon > 0$ there is a neighborhood U of x_0 such that

$$f(x, r) - a_\alpha r^\alpha + \epsilon \mid r \mid \geqq 0 \text{ for } (x, r) \in R, \qquad x \in U.$$

Then $g(x, r) - a_\alpha r^\alpha + \epsilon \mid r \mid \geqq 0$ for $x \in U$ and all r; and so the lower limit function of $g - a_\alpha r^\alpha + \epsilon \mid r \mid$, which is $F(x, r) - a_\alpha r^\alpha + \epsilon \mid r \mid$, is also non-negative in U. In particular, $F(x_0, r) \geqq a_\alpha r^\alpha - \epsilon \mid r \mid$. Since ϵ is arbitrary, we have $F(x_0, r) \geqq a_\alpha r^\alpha$.

LEMMA 3.4. *For every $M > 0$ there exists a sequence of functions $g_n(x, r)$, defined and continuous in (x, r) for $\mid x \mid \leqq M$, $\mid r \mid \leqq M$, convex in r for fixed x, and such that*

(i) $$g_1(x, r) < g_2(x, r) < \cdots$$

for $\mid x \mid \leqq M$, $\mid r \mid \leqq M$, and

(ii) $$\lim_{n\to\infty} g_n(x, r) = F(x, r), \qquad (x, r) \in R, \mid x \mid \leqq M, \qquad \mid r \mid \leqq M.$$

Proof. On the bounded closed set $\mid x \mid \leqq M$, $\mid r \mid \leqq M$ the l.s.c. function F attains its lower bound. Since $F \not\equiv -\infty$, this lower bound is not $-\infty$. Consequently,[10] there exists a sequence $\{\phi_n(x, r)\}$ of functions continuous for $\mid x \mid \leqq M$, $\mid r \mid \leqq M$, such that $\phi_1(x, r) < \phi_2(x, r) < \cdots \to F(x, r)$. For each x let $g_n(x, r)$ be the "convex envelope" of $\phi_n(x, r)$; that is, the least upper bound of all convex functions $\psi(r) \leqq \phi_n(x, r)$. Then $g_n(x, r)$ is convex in r. From the definition it is easily seen that if two functions differ by less than ϵ, then so do their convex envelopes. As $\bar{x} \to x$, the function $\phi_n(\bar{x}, r)$ tends to $\phi_n(x, r)$ uniformly in r, so $g_n(\bar{x}, r)$ tends to $g_n(x, r)$ uniformly in r. For each fixed x, $g_n(x, r)$ is convex in r, hence continuous in r. Thus $g_n(x, r)$ is continuous in r and is continuous in x uniformly with respect to r, so it is continuous in both variables. From $\phi_1 < \phi_2 < \cdots$ it follows at once that $g_1 < g_2 < \cdots$.

Finally, suppose $(x_0, r_0) \in R$, $\mid x \mid \leqq M$, $\mid r \mid \leqq M$, and let u be any number less than $F(x_0, r_0)$. By Lemma 3.3 there is a linear function $a_\alpha r^\alpha$ such that $a_\alpha r_0^\alpha > u$ and $a_\alpha r^\alpha \leqq F(x_0, r)$ for all r. If we put $a_0 = \frac{1}{2}(u - a_\alpha r_0^\alpha) < 0$, then $a_0 + a_\alpha r_0^\alpha > u$ and $F(x_0, r) > a_0 + a_\alpha r^\alpha$ for all r. The continuous functions $\phi_n(x_0, r) - a_0 - a_\alpha r^\alpha$ tend on the bounded closed set $\mid r \mid \leqq M$ to the positive limit $F(x_0, r) - a_0 - a_\alpha r^\alpha$; hence for all large n we have $\phi_n(x_0, r) > a_0 + a_\alpha r^\alpha$ for all r with $\mid r \mid \leqq M$. Then $a_0 + a_\alpha r^\alpha$ is a convex function which does not exceed $\phi_n(x_0, r)$, so for the least upper bound g_n of such functions we have $g_n(x_0, r) \geqq a_0 + a_\alpha r^\alpha$. In particular, $g_n(x_0, r_0) \geqq a_0 + a_\alpha r_0^\alpha > u$; so lim $g_n(x_0, r_0) \geqq u$. This being true for all $u < F(x_0, r_0)$, it follows that lim $g_n(x_0, r_0) \geqq F(x_0, r_0)$. On the other hand, $g_n(x_0, r_0) \leqq F(x_0, r_0)$ for all n, so lim $g_n(x_0, r_0) \leqq F(x_0, r_0)$. Therefore $g_n(x_0, r_0)$ tends to $F(x_0, r_0)$. This establishes the lemma.

[10] Carathéodory, *Vorlesungen über Reelle Funktionen*, p. 402.

LEMMA 3.5. *In Lemma 3.4 we can further assume that the first partial derivatives of the g_n with respect to the r^i exist and are continuous for $|x| \leqq M$, $|r| \leqq M$.*

Proof. In the statement of Lemma 3.4 we replace M by $M + 1$ and denote the functions thus obtained by $f_n(x, r)$. If $0 < h < 1/(q + 1)$, the integral

$$(3.1)\qquad \begin{aligned} g_n^{(h)}(x, r) &= \frac{1}{(2h)^{q+1}} \int_{r^0-h}^{r^0+h} \cdots \int_{r^q-h}^{r^q+h} f_n(x^0, \cdots, x^q, t^0, \cdots, t^q)dt^0 \cdots dt^q \\ &= \frac{1}{(2h)^{q+1}} \int_{-h}^{h} \cdots \int_{-h}^{h} f(x^0, \cdots, x^q, r^0 + t^0, \cdots, r^q + t^q)dt^0 \cdots dt^q \end{aligned}$$

is defined for $|x| \leqq M$, $|r| \leqq M$. Since f_n is continuous for $|x| \leqq M + 1$, $|r| \leqq M + 1$, the integral is a continuous function of x and r. It is a convex function, as we see by integrating the inequality

$$\{f_n(x, r_1 + t) + f_n(x, r_2 + t)\}/2 \geqq f_n(x, \tfrac{1}{2}(r_1 + r_2) + t).$$

It has continuous first partial derivatives with respect to the r^i.

Since $f_{n-1} < f_n < f_{n+1}$, by Lemma 3.4, there is a positive number ϵ_n such that $f_{n+1}(x, r) - f_n(x, r) > \epsilon_n$ and $f_n(x, r) - f_{n-1}(x, r) > \epsilon_n$ for $|x| \leqq M + 1$ and $|r| \leqq M + 1$. Also, f_n is continuous, so there exists a $\delta > 0$ such that $|f_n(x, \bar{r}) - f_n(x, r)| < \epsilon_n$ if $|\bar{r} - r| \leqq \delta$. Choosing $h = \delta$, we have from the definition (3.1)

$$f_{n-1}(x, r) < f_n(x, r) - \epsilon_n < g_n^{(\delta)}(x, r) < f_n(x, r) + \epsilon_n < f_{n+1}(x, r).$$

Thus $f_1 < g_2^{(\delta)} < f_3 < g_4^{(\delta)} < \cdots$. Choosing the second, fourth, $\cdots$ terms of this sequence and renaming them $g_1, g_2, \cdots$, we have $g_1 < g_2 < \cdots$, and also $\lim g_n(x, r) = \lim f_n(x, r) = F(x, r)$ if $(x, r) \epsilon R$ and $|x| \leqq M$, $|r| \leqq M$.

LEMMA 3.6. *For each of the functions $g_n(x, r)$ of Lemma 3.5 the integral $\int g_n(x, \dot{x})dt$ is l.s.c. on the class of all functions $x(t)$, $0 \leqq t \leqq 1$, satisfying the Lipschitz condition of constant M and the condition $|x(t)| \leqq M$.*

Proof. Suppose that the functions $x_n(t)$ satisfy the above conditions and converge uniformly to $x_0(t)$. Then $x_0(t)$ also satisfies these conditions. Let $g(x, r)$ be any one of the functions $g_n(x, r)$, and let $g_{(i)}(x, r)$ be the partial derivative of g with respect to r^i. Since g is uniformly continuous for $|x| \leqq M$ and $|r| \leqq M$, the difference $g(x_n(t), r) - g(x_0(t), r)$ tends to zero uniformly for $0 \leqq t \leqq 1$ and $|r| \leqq M$. Therefore

$$(3.2)\qquad \liminf \int_0^1 g(x_n, \dot{x}_n)dt = \liminf \int_0^1 g(x_0, \dot{x}_n)dt.$$

For fixed x, the linear function osculating $g(x, r)$ at r_0 is

$$g(x, r_0) + (r^\alpha - r_0^\alpha)\, g_{(\alpha)}(x, r_0).$$

Since g is convex in r, it is not less than this linear function:

$$g(x, r) \geqq g(x, r_0) + (r^\alpha - r_0^\alpha)g_{(\alpha)}(x, r_0).$$

Therefore

$$\text{(3.3)}\qquad \begin{aligned}\liminf \int_0^1 g(x_0, \dot{x}_n)dt \geqq \int_0^1 g(x_0, \dot{x}_0)dt \\ + \liminf \int_0^1 (\dot{x}_n^\alpha - \dot{x}_0^\alpha)\, g_{(\alpha)}(x_0, \dot{x}_0)dt.\end{aligned}$$

From $x_n \rightrightarrows x_0$ we find that for $0 \leqq h \leqq k \leqq 1$

$$\lim_{n\to\infty} \int_h^k (\dot{x}_n^i - \dot{x}_0^i)dt = \lim_{n\to\infty} [x_n^i(k) - x_0^i(k)] - [x_n^i(h) - x_0^i(h)] = 0.$$

Also, $|\dot{x}_n^i - \dot{x}_0^i| \leqq 2M$. Hence[11] for each i we have

$$\lim_{n\to\infty} \int_0^1 (\dot{x}_n^i - \dot{x}_0^i)\, g_{(i)}(x_0, \dot{x}_0)dt = 0.$$

Therefore the sum on the right in (3.3) tends to zero, and by (3.2) and (3.3) we have

$$\liminf \int_0^1 g(x_n, \dot{x}_n)dt \geqq \int_0^1 g(x_0, \dot{x}_0)dt.$$

This establishes the lemma.

We now take up the proof of Theorem 3.1. Let $[M]$ be the class of admissible functions $x(t)$, $0 \leqq t \leqq 1$ such that $|x(t)| \leqq M$ and $|\dot{x}(t)| \leqq M$. With the g_n of Lemma 3.5 we have for each $x(t) \in M$

$$\int_0^1 g_n(x, \dot{x})dt \leqq \int_0^1 g_{n+1}(x, \dot{x})dt.$$

For almost all t the set $(x(t), \dot{x}(t)) \in R$, and for all such t we know by Lemma 3.5 that $g_n(x(t), \dot{x}(t))$ increases with n and tends to $F(x(t), \dot{x}(t))$. So

$$\int g_n(x, \dot{x})dt \to \int F(x, \dot{x})dt,$$

and on $[M]$ the functional $\int F(x, \dot{x})dt$ is the limit of an increasing sequence of functionals $\int g_n(x, \dot{x})dt$. By Lemma 3.6, these last are l.s.c., so $\int F(x, \dot{x})dt$ is itself l.s.c. on $[M]$. But on all admissible curves, and in particular on $[M]$, we have $\int F(x, \dot{x})dt = \int f(x, \dot{x})dt$, by Lemma 3.3, so that $\int f(x, \dot{x})dt$ is l.s.c. on $[M]$. By Lemma 3.2, this implies that $\mathfrak{F}(C)$ is l.s.c.

§4. Theorem 3.1 appears to apply only to integrals in parametric form. But, as a matter of fact, the hypotheses are weak enough so that we can state a theorem exactly equivalent to Theorem 3.1 in which the notation is that of ordinary problems. We wish then to investigate the semi-continuity of integrals $\int g(u, y, y')\, du$ on classes of functions $y = y(u)$, $a \leqq u \leqq b$. But we cannot even define semi-continuity until we have a notion of limit defined. Accordingly, if $y = y(u)$, $a \leqq u \leqq b$, and $y = y_1(u_1)$, $a_1 \leqq u_1 \leqq b_1$, are continuous

[11] Hobson, *Theory of Functions of a Real Variable*, vol. I, §279.

functions, we define the distance dist (y, y_1) as follows. First we extend the range of y to the whole u-axis by setting $y(u) = y(a)$ for $u < a$ and $y(u) = y(b)$ for $u > b$, and we extend the range of y_1 similarly. We then define the dist (y, y_1) to be the greatest of the three numbers max $| y(u) - y_1(u) |$, $| a - a_1 |$, $| b - b_1 |$. The distance thus defined actually does define a metric for continuous functions, but what we are interested in showing is that if the curves C_n are defined by the continuous functions $y = y_n(u)$, $a_n \leqq u \leqq b_n$, $(n = 0, 1, \cdots)$, and dist $(y_0, y_n) \to 0$, then $\lim C_n = C_0$. Let us map the interval (a_0, b_0) on (a_n, b_n) by a linear transformation $u_n(u)$; then the maximum of $| u_n(u) - u |$ occurs at one of the ends of (a_0, b_0), and is either $| a_n - a_0 |$ or $| b_n - b_0 |$. In any case max $| u_n(u) - u | \leqq$ dist $(y_n, y_0) \to 0$. Now write

$$\begin{aligned} \text{dist } (C_n, C_0) &\leqq \max | y_n(u_n(u)) - y_0(u) |, \qquad a_0 \leqq u \leqq b_0 \\ &\leqq \max | y_n(u_n(u)) - y_0(u_n(u)) | + \max | y_0(u_n(u)) - y_0(u) |. \end{aligned}$$

By definition, the first term on the right does not exceed dist $(y_n, y_0) \to 0$. Since $y_0(u)$ is uniformly continuous and $u_n(u) \rightrightarrows u$, the second term also tends to zero, and so $C_n \to C_0$.

THEOREM 4.1. *Let $g(u, y, \eta)$ satisfy the following conditions:*

(4.1a) *$g(u, y, \eta)$ is defined (finite or $+\infty$) and l.s.c. on a set Y in (u, y, η)-space;*

(4.1b) *the set of (u, y) such that $(u, y, \eta) \in Y$ for some η is closed;*

(4.1c) *if $(u_0, y_0, \eta_0) \in Y$ and $h < g(u_0, y_0, \eta_0)$, there exists a linear function $a_0 + a_\alpha r^\alpha$ such that* (i) *$a_0 + a_\alpha \eta^\alpha > h$ and* (ii) *for every $\epsilon > 0$ there is a neighborhood U of (u_0, y_0) such that $g(u, y, \eta) \geqq a_0 + a_\alpha \eta^\alpha - \epsilon(1 + | \eta |^2)^{\frac{1}{2}}$ if $(u, y, \eta) \in Y$ and $(u, y) \in U$.*

Then for every $M > 0$ the integral $\mathcal{G}(y) = \int g(u, y, \dot{y})du$ is l.s.c. on the class of all functions $y = y(u)$, $a \leqq u \leqq b$, having total variation $\leqq M$ and such that $(u, y(u), \dot{y}(u)) \in Y$ for almost all u.

It is not difficult to see that this theorem implies Theorem 3.1. If conditions (2.1) are satisfied, we introduce the new notation $(y^1, \cdots, y^{q+1}) = (x^0, \cdots, x^q)$, $(\eta^1, \cdots, \eta^{q+1}) = (r^0, \cdots, r^q)$, and define $g(u, y, \eta) = f(x, r)$. The set Y will consist of all (u, y, η) for which $(y, \eta) \equiv (x, r)$ is in R. Then (4.1a) and (4.1c) follow from (2.1a) and (2.1d), respectively. By (2.1b), if $(u, y, \eta) \in Y$ for some η, so is $(u, y, 0)$, so that (2.1c) implies (4.1b). Now let C_n: $x = \bar{x}_n(t)$, $a_n \leqq t \leqq b_n$, be a sequence of admissible curves of length $\leqq M$ tending to C_0. For each C_j $(j = 0, 1, \cdots)$ we can choose an a.c. representation $x = x_j(u)$, $0 \leqq u \leqq 1$, so that $x_n \rightrightarrows x_0$. By Theorem 4.1 we have

$$\begin{aligned} \liminf \int_0^1 f(x_n, \dot{x}_n)du &= \liminf \int_0^1 g(u, y_n, \dot{y}_n)du \\ &\geqq \int_0^1 g(u_0, y_0, \dot{y}_0)du \\ &= \int_0^1 f(x_0, \dot{x}_0)du. \end{aligned}$$

This establishes Theorem 3.1.

On the other hand, Theorem 3.1 implies Theorem 4.1. Suppose conditions (4.1) verified. We define

$$f(u, y, \xi, \eta) = \xi g(u, y, \eta/\xi) \text{ for } \xi > 0,$$

$$f(u, y, 0, 0) = 0,$$

and we define R in the following way: if $(u, y, \eta) \in Y$, then $(u, y, t, t\eta) \in R$ for all $t \geqq 0$. We now introduce the notation $(x^0, x^1, \cdots, x^q) = (u, y^1, \cdots, y^q)$, $(r^0, r^1, \cdots, r^q) = (\xi, \eta^1, \cdots, \eta^q)$. Then condition (2.1b) is satisfied. The set of x for which $(x, 0) \in R$ is the same as the set (u, y) for which $(u, y, \eta) \in Y$ for some η, and is closed by (4.1b), so that the second part of (2.1c) holds. The first part of (2.1c) follows from the definition of R, since R consists of the points $(u, y, t, t\eta)$, each of which belongs to a line-segment of points of R.

Let $(x_0, r_0) \equiv (u_0, y_0, t, t\eta)$, $t \geqq 0$, be a point of R, and let h be any number less than $f(x_0, r_0)$. Then if $t > 0$,

$$ht^{-1} < t^{-1}f(x_0, r_0) = f(u_0, y_0, 1, \eta) = g(u_0, y_0, \eta_0).$$

By (4.1c) there is a function $a_0 + a_\alpha \eta^\alpha$ such that (i) $a_0 + a_\alpha \eta^\alpha > ht^{-1}$, and (ii) for $\epsilon > 0$ there exists a neighborhood U of (u_0, y_0) such that

$$g(u, y, \eta) \geqq a_0 + a_\alpha \eta^\alpha - \epsilon(1 + |\eta|^2)^{\frac{1}{2}}$$

for $(u, y, \eta) \in Y$, $(u, y) \in U$. If we write f for g and use the homogeneity of f, these become

(i) $$a_\alpha r^\alpha = a_0 t + a_\beta(t\eta^\beta) > h,$$

(ii) $$f(x, r) = f(u, y, t, t\eta) \geqq a_0 t + a_\beta(t\eta^\beta) - \epsilon(t^2 + |t\eta|^2)^{\frac{1}{2}}$$
$$= a_\alpha r^\alpha - \epsilon |r|,$$

as required in (2.1d). If, however, $t = 0$, we notice that the discussion of (ii) requires no alteration, while (i) reduces to $a_\alpha r_0^\alpha = 0 = f(x, r_0) > h$. Therefore (2.1d) holds.

To establish the fact that $f(x, r)$ is l.s.c. on R, we first consider a set (x_0, r_0) with $r_0 = 0$. If $N - 1$ is the length of the vector a_i of (2.1d), and if we there take $\epsilon = 1$, we find that for every $\delta > 0$ we have

$$f(x, r) \geqq a_\alpha r^\alpha - |r| \geqq -N|r| > -\delta = f(x_0, r_0) - \delta,$$

provided that $(x, r) \in R$, $x \in U$, $|r| < \delta/N$. So $f(x, r)$ is l.s.c. at (x_0, r_0). If $|r_0| \neq 0$, then $r_0^0 > 0$, for since $(x_0, r_0) = (u_0, y_0, t, t\eta_0)$, the only way to have $r_0^0 = 0$ is to have $t = 0$. In this case, if we choose $(x_n, r_n) \equiv (u_n, y_n, \xi_n, \eta_n)$ tending to (x_0, r_0), we have $\xi_n > 0$ for almost all n, and

$$\liminf f(x_n, r_n) = \liminf \xi_n g(u_n, y_n, \eta_n/\xi_n)$$
$$\geqq \xi_0 g(u_0, y_0, \eta_0/\xi_0) = f(x_0, r_0).$$

So in this case also $f(x, r)$ is l.s.c. at (x_0, r_0), and (2.1a) is satisfied.

Suppose now that $y_n(u)$, $a_n \leqq u \leqq b_n$, is a sequence of admissible[12] functions tending to $y_0(u)$, $a_0 \leqq u \leqq b_0$. If we define C_j to be the curve $x^0 = u$, $x^i = y^i_j(u)$, $a_j \leqq u \leqq b_j$, then $C_n \to C_0$, and by Theorem 3.1

$$\begin{aligned}\liminf \int_{a_n}^{b_n} g(u, y_n, \dot{y}_n)du &= \liminf \int_{a_n}^{b_n} f(x_n, \dot{x}_n)du \\ &\geqq \int_{a_0}^{b_0} f(x_0, \dot{x}_0)du \\ &= \int_{a_0}^{b_0} g(u, y_0, \dot{y}_0)du.\end{aligned}$$

This establishes Theorem 4.1.

§5. We now begin to deduce from Theorems 3.1 and 4.1 corollaries of more recognizable appearance. An immediate corollary of Theorem 3.1 is

THEOREM 5.1. *If $f(x, r)$ is defined and l.s.c. for all x in a closed set A and all r, and $f(x, tr) = tf(x, r)$ for $x \in A$ and $t \geqq 0$, and $f(x, r)$ is a convex function of r for each $x \in A$, then for every $M > 0$ the integral $\mathfrak{F}(C)$ is l.s.c. on the class of all curves lying in A and having length $\leqq M$.*

A second corollary is

THEOREM 5.2. *Let the functions $f(x, r)$ and $\phi^k(x, r)$ $(k = 1, 2, \cdots, m)$ be defined and continuous for all x in a perfect set A and all r; let the partial derivatives of f and ϕ^k with respect to the r^i exist and be continuous for $x \in A$ and $|r| > 0$; let $f(x, tr) = tf(x, r)$ and $\phi^k(x, tr) = t\phi^k(x, r)$ for $x \in A$ and $t \geqq 0$; for each $x \in A$, let there exist constants $c_1, \cdots, c_m$ such that the function*

$$F(x, r) \equiv f(x, r) + c_\alpha \phi^\alpha(x, r)$$

satisfies the inequality

$$\mathcal{E}_F(x, r, \bar{r}) \equiv F(x, \bar{r}) - \bar{r}^\alpha F_{(\alpha)}(x, r) \geqq 0$$

whenever $\phi^k(x, r) = \phi^k(x, \bar{r}) = 0$ and $|r| > 0$. Then for every $M > 0$ the integral $\mathfrak{F}(C) = \int f\, dt$ is l.s.c. on the class of all curves $x = x(t)$ lying in A, having length $\leqq M$, and satisfying the equations $\phi^k(x(t), \dot{x}(t)) = 0$ for almost all t.

Let R be the class of all sets (x, r) such that $x \in A$ and $\phi^k(x, r) = 0$, $(k = 1, \cdots, m)$; this set is closed. Conditions (2.2a, b, c) are clearly satisfied. If $(x_0, r_0) \in R$ and $|r_0| \neq 0$, we set $a_\alpha = F_{(\alpha)}(x_0, r_0)$. It is a well-known consequence of the homogeneity of F that

$$r_0^\alpha F_{(\alpha)}(x_0, r_0) = F(x_0, r_0), \tag{5.3}$$

while by hypothesis for all r such that $(x_0, r) \in R$ we have

$$a_\alpha r^\alpha = r^\alpha F_{(\alpha)}(x_0, r_0) = F(x_0, r) - \mathcal{E}_F(x_0, r_0, r) \leqq F(x_0, r).$$

[12] A function $y(u)$ is admissible if $(u, y, \dot{y}) \in Y$ for almost all u.

But if $(x, r) \epsilon R$, by definition all ϕ^k vanish and $F(x, r) = f(x, r)$. Hence

$$a_\alpha r_0^\alpha = f(x_0, r_0) \text{ and } a_\alpha r^\alpha \leqq f(x_0, r) \text{ for } (x_0, r) \epsilon R.$$

Thus in case $|r_0| > 0$, condition (2.2d) holds.

If $r_0 = 0$, we distinguish two cases. There may be no r_0 except 0 for which $\phi^k(x_0, r) = 0$. In this case we choose numbers a_i arbitrarily and obtain $a_\alpha r^\alpha = 0 = f(x_0, r)$ for all r such that $(x_0, r) \epsilon R$, namely r_0. Or there may be an $r_1 \neq 0$ such that $(x_0, r_1) \epsilon R$. In this case there exists, as we have seen above, a function $a_\alpha r^\alpha$ such that $a_\alpha r^\alpha \leqq f(x_0, r)$ for $(x_0, r) \epsilon R$, while $a_\alpha r_0^\alpha = 0 = f(x_0, r_0)$. So in any case (2.2d) is satisfied, and our conclusion holds.

§6. Just as Theorem 3.1 led us to a theorem on Lagrange problems in parametric form, so does Theorem 4.1 lead us to one on Lagrange problems in ordinary form.

THEOREM 6.1. *Let the functions $g(u, y, \eta)$ and $\psi^k(u, y, \eta)$, $(k = 1, \cdots, m)$, satisfy the conditions*

(6.1) *$g(u, y, \eta)$ and the $\psi^k(u, y, \eta)$ are continuous and possess continuous first partial derivatives with respect to the η^i for all (u, y) in a closed set A and all η;*

(6.2) *for each $(u, y) \epsilon A$ the system of equations $\psi^k(u, y, \eta) = 0$ has at least one solution;*

(6.3) *there exists a set of functions $c_i(u, y)$, $(i = 1, \cdots, m)$, continuous on A, such that if we set $G = g + c_\alpha \psi^\alpha$, for every admissible set (u_0, y_0, η_0) and every $\epsilon > 0$ there is a neighborhood U of (u_0, y_0) for which*

$$G(u, y, \eta) - G(u_0, y_0, \eta_0) - (\eta^\alpha - \eta_0^\alpha)G_{(\alpha)}(u_0, y_0, \eta_0) \geqq -\epsilon(1 + |\eta|^2)^{\frac{1}{2}}$$

whenever (u, y, η) is admissible and $(u, y) \epsilon AU$.

Then for every $M > 0$ the integral $\mathcal{G}(y) = \int g(u, y, \dot{y})du$ is l.s.c. on the class of all admissible functions $y = y(u)$, $a \leqq u \leqq b$, having total variation $\leqq M$ and such that $(u, y(u)) \epsilon A$ for every u.

Referring to Theorem 4.1, we observe that conditions (4.1, a, b) are obviously fulfilled, while (6.3) implies (4.1c) if we set

$$a_0 = G(u_0, y_0, \eta_0) - \eta_0^\alpha G_{(\alpha)}(u_0, y_0, \eta_0)$$

$$a_i = G_{(i)}(u_0, y_0, \eta_0) \qquad (i = 1, \cdots, q)$$

and recall that

$$G(u, y, \eta) = g(u, y, \eta) + c_\alpha \psi^\alpha(u, y, \eta) = g(u, y, \eta)$$

for admissible (u, y, η).

From this theorem there follows a rather interesting corollary.

THEOREM 6.2. *Suppose that*

(6.4) *$g(u, y, \eta)$ and $\psi^k(u, y, \eta)$ are defined and continuous, together with their first partial derivatives with respect to the η^i, for all (u, y) in a closed set A and all η;*

(6.5) *for each $(u, y) \epsilon A$ the system of equations $\psi^k(u, y, \eta) = 0$ has at least one solution;*

(6.6) *for every admissible set* (u_0, y_0, η_0) *and every* $\delta > 0$ *there is a neighborhood* U *of* (u_0, y_0) *such that if* (u, y) *is in* AU, *the equations* $\psi^k(u, y, \eta) = 0$ *have a solution*[13] *with* $|\eta - \eta_0| < \delta$;
(6.7) *there exist functions* $c_i(u, y)$ *continuous on* A *such that if we set* $G = g + c_\alpha \psi^\alpha$, *then the inequality*

$$\mathcal{E}_G(u, y, \eta, \bar{\eta}) \equiv G(u, y, \bar{\eta}) - G(u, y, \eta) - (\bar{\eta}^\alpha - \eta^\alpha)G_{(\alpha)}(u, y, \eta) \geqq 0$$

holds for all admissible sets (u, y, η) *and* $(u, y, \bar{\eta})$.
Then for every $M > 0$ *the integral* $\mathcal{G}(y) \equiv \int g(u, y, \dot{y})du$ *is l.s.c. on the class of all a.c. functions* $y = y(u)$ *satisfying the equations* $\psi^k(u, y, \dot{y}) = 0$ *almost everywhere and such that* $(u, y(u)) \in A$ *and having total variation of* $y(u) \leqq M$.

Comparing this with Theorem 6.1, we see that the only hypothesis not obviously satisfied is (6.3). The functions G and $G_{(i)}$ are continuous for all $(u, y) \in A$ and all η; hence if (u_0, y_0, η_0) is admissible, for every $\gamma > 0$ there is a neighborhood U_1 of (u_0, y_0) and a $\delta > 0$ such that if $(u, y) \in AU_1$ and $|\eta - \eta_0| < \delta$, then

$$(6.8) \qquad |G(u_0, y_0, \eta_0) - G(u, y, \eta)| < \gamma,$$
$$\{\Sigma_\alpha[G_{(\alpha)}(u_0, y_0, \eta_0) - G_{(\alpha)}(u, y, \eta)]^2\}^{\frac{1}{2}} < \gamma.$$

Let U be the neighborhood mentioned in (6.6). For every (u, y) in AU_1U there is an $\bar{\eta}$ with $|\bar{\eta} - \eta_0| < \delta$ for which $(u, y, \bar{\eta})$ is admissible. Then if (u, y, η) is admissible and $(u, y) \in AU_1U$, we have

$$\begin{aligned} G(u, y, \eta) &- G(u_0, y_0, \eta_0) - (\eta^\alpha - \eta_0^\alpha)G_{(\alpha)}(u_0, y_0, \eta_0) \\ &= [G(u, y, \eta) - G(u, y, \bar{\eta}) - (\eta^\alpha - \bar{\eta}^\alpha)G_{(\alpha)}(u, y, \bar{\eta})] \\ &+ [G(u, y, \bar{\eta}) - G(u_0, y_0, \eta_0)] + (\eta^\alpha - \bar{\eta}^\alpha)[G_{(\alpha)}(u, y, \bar{\eta}) - G_{(\alpha)}(u_0, y_0, \eta_0)] \\ &+ (\eta_0^\alpha - \bar{\eta}^\alpha)G_{(\alpha)}(u_0, y_0, \eta_0). \end{aligned}$$

The first term on the right is non-negative by (6.7). The second is not less than $-\gamma$. The third is not less (by 6.8) than $-(|\eta| + |\bar{\eta}|)\cdot\gamma$. If we write P for the length of the vector whose components are $G_{(i)}(u_0, y_0, \eta_0)$, the fourth term is not less than $|\bar{\eta} - \eta_0|P < \cdot P\delta$, so the left member is greater than $-\gamma - \gamma|\eta| - \gamma|\bar{\eta}| - P\delta$. For any $\epsilon > 0$ we can choose δ and γ so small that $\gamma(1 + |\eta_0| + \delta) + P\delta < \epsilon/2$; then the left member is greater than $-\epsilon/2 - \epsilon|\eta|/2 > -\epsilon[1 + |\eta|^2]^{\frac{1}{2}}$. This proves that (6.3) is satisfied (with U replaced by U_1U), and so the conclusion of Theorem 6.2 must hold.

§7. If we compare Theorem 6.2 with Theorem 5.2 we notice that there is a decided strengthening of hypotheses. Hypotheses (6.5) and (6.6) have no analogues in Theorem 5.2, and even (6.7) requires that the c_i be continuous, which was not needed in §5. This suggests an investigation to see whether these hypotheses are really essential or are merely dictated by our methods of

[13] This holds in particular if $m < q$ and the matrix $\|\psi^k_{(i)}(u, y, \eta)\|$ has rank m for all admissible (u, y, η).

proof. We shall here show by examples that the former is the case; if either (6.5) or (6.6) is omitted, or even if (6.7) is relaxed to allow discontinuous $c_i(u, y)$, the theorem is no longer valid.

For all $y \neq 0, \pm 1, \pm\frac{1}{2}, \pm\frac{1}{3}, \cdots$ we define $\mu(y) = e^{-y^{-4}\sin^{-2}(\pi/y)}$, while for these exceptional values we set $\mu(y) = 0$. Then $\mu(y)$ is defined and continuous for all y, and is positive except for $y = 0, \pm 1, \cdots, \pm 1/n, \cdots$. Our integrand will be $g(u, y, \eta) = -\eta^2 y^2 \sin^2(\pi/y)$, if $y \neq 0$, $g(u, 0, \eta) \equiv 0$. For our three examples we choose three different side equations $\psi(u, y, \eta) = 0$.

Example I: $\psi(u, y, \eta) = e^\eta - \mu(y)$.

Example II: $\psi(u, y, \eta) = q(y, \mu(y) + 1 - e^\eta)$, where $q(y, v) = v^3 - v^2 + \mu(y)(v - 1)$.

Example III: $\psi(u, y, \eta) = (e^\eta - \mu(y))(e^\eta - 1)$.

In example II, we observe that $q(y, v) = 0$ has one real root $v = 1$ if $\mu(y) > 0$, and two distinct roots $v = 0, v = 1$ if $\mu(y) = 0$. Hence in this case the solution of the equation $\psi(u, y, \eta) = 0$ is $\eta = \log \mu(y)$ if $y \neq 0, \pm 1, \pm 1/2, \cdots$, and $\eta = 0$ if $y = 0, \pm 1, \cdots$. In example I, for $y \neq 0, \pm 1, \cdots$ the solution is $\eta = \log \mu(y)$, and for $y = 0, \pm 1, \pm 1/2, \cdots$ there is no solution. In example III, $\eta = \log \mu(y)$ is a solution if $y \neq 0, \pm 1, \pm 1/2, \cdots$, and $\eta = 0$ is a solution for all y.

In examples I and II hypothesis 6.7 is satisfied if, for example, we take $c \equiv 0$. For, given any (u, y), there is never more than one η such that (u, y, η) is admissible, and (6.7) reduces to the triviality $\mathcal{E}_G(u, y, \eta, \eta) = 0$. Example I satisfies (6.6)[14] but not (6.5); example II satisfies (6.5), but not (6.6). Example III satisfies both of these hypotheses. With regard to (6.7), we notice that if $y = 0, \pm 1, \cdots$, there is only one η (namely 0) for which $\psi(u, y, \eta) = 0$, so as before, if (u, y, η) and $(u, y, \bar\eta)$ are admissible, $\mathcal{E}_G(u, y, \eta, \bar\eta) = \mathcal{E}_G(u, y, 0, 0) = 0$, no matter how we choose c. If $y \neq 0, \pm 1, \cdots$, there are two solutions of $\psi = 0$, namely 0 and $\log \mu(y)$. Then $\psi_\eta(u, y, 0) = 1 - \mu(y) > 0$, while $\psi_\eta(u, y, \log \mu(y)) = \mu(y)(\mu(y) - 1) < 0$. Let us set $G = g + c\psi$ and try to determine c so that $\mathcal{E}_G(u, y, \eta, \bar\eta) \geqq 0$ whenever (u, y, η) and $(u, y, \bar\eta)$ are both admissible. If $\eta = \bar\eta$ this is certainly satisfied. Otherwise we have

$$\mathcal{E}_G(u, y, 0, \log \mu(y)) = -y^{-6}\sin^{-4}(\pi/y) + y^{-4}\sin^{-2}(\pi/y)[c(1 - \mu(y))],$$

$$\mathcal{E}_G(u, y, \log \mu(y), 0) = y^{-6}\sin^{-4}(\pi/y) - y^{-4}\sin^{-2}(\pi/y)[2y^{-2} + c\mu(y)(\mu(y) - 1)].$$

These are both positive if c is large enough. Hence (6.7) holds except for the requirement that the c_i be continuous.

In all these examples there is a family of admissible functions determined by the equation

$$\dot y = \log \mu(y) = -y^{-4}\sin^{-2}(\pi/y); \tag{7.1}$$

these are absolutely continuous and monotonic decreasing, with derivatives less than -1 for $|y| \leqq 1$. For example II there is a second family of admissible

[14] In fact, if $\psi(u, y, \eta) = e^\eta - 1 = 0$, then $\partial\psi/\partial\eta = e^\eta = 1$.

functions $y = 0$, $y = \pm 1$, $y = \pm\frac{1}{2}$, $\cdots$. For example III every function $y =$ constant is admissible.

To establish the lack of semi-continuity in example I, we let y_0 be the function defined only for $a = 0$, and there having the value 0, and we define y_n to be a solution of (7.1) in the interval $(0, b_n)$, where $y_n(0) = 1/n$ and b_n is so chosen that $y_n(b_n) = 1/(n + 1)$. Then $y_n \to y_0$, and $\mathcal{G}(y_0) = 0$. We calculate $\mathcal{G}(y_n)$ most easily by using y as the independent variable:

$$\mathcal{G}(y_n) = \int_0^{b_n} - y^2 \sin^2(\pi/y)\dot{y}^2 dx = \int_{1/(n+1)}^{1/n} - y^2 \sin^2(\pi/y) y^{-4} \sin^{-2}(\pi/y) dy$$
$$= [y^{-1}]_{1/(n+1)}^{1/n} = -1.$$

So example I is not l.s.c., even though the total variation of y_n is $1/n - 1/(n + 1)$, which tends to 0.

The same family of functions is admissible for examples II and III also, and so they are also not l.s.c. However, it is interesting to modify the y_n somewhat. We define $y_0(u) \equiv 0$, $0 \leqq u \leqq 1$; and for $n = 1, 2, \cdots$ we define $y_n(u)$ to be the solution of (7.1) on $(0, b_n)$ as before, setting $y_n(u) \equiv 1/(n + 1)$ for $b_n < u \leqq 1$. The functions y_0, y_n thus defined are absolutely continuous on the interval $(0, 1)$, and since on $(b_n, 1)$ we have $g(u, y_n, \dot{y}_n) = g(u, y_n, 0) = 0$, we still have $\mathcal{G}(y_n) = -1$. For example III we thus see that the function $y_0 = 0$, $0 \leqq u \leqq 1$, furnishes a weak relative minimum for $\mathcal{G}(y)$, since if we consider only admissible functions $y(u)$ with $|y| \leqq 1$ and $|\dot{y}| \leqq 1$ we obtain only the functions $y =$ const., for which $\mathcal{G}(y) = 0$. Of course there is not a strong relative minimum at y_0, as our comparison functions y_n show. The function y_0 is imbedded in the field of extremals $y =$ const. The Weierstrass condition holds along y_0, but not in strengthened form.

In the usual Lagrange problem the number m of equations is required to be less than the number q of functions $y(u)$, while here $m = q = 1$. This objection is readily disposed of if we interpret our problem as one in (x, y, z)-space in which the functions g and ψ happen to be independent of z and $\dot{z}$.

§8. From §5 we obtain at once a theorem concerning integrals in ordinary form, without side equations; for if we suppose that there are no functions ψ^k, Theorem 6.2 becomes

THEOREM 8.1. *If $g(u, y, \eta)$ is defined and continuous, together with its first partial derivatives $g_{(i)}$, for all (u, y) in a closed set A and all η, and if $\mathcal{E}(u, y, \eta, \bar{\eta}) \geqq 0$ for all $(u, y) \epsilon A$ and all η and $\bar{\eta}$, then for every $M > 0$ the integral $\mathcal{G}(y)$ is l.s.c. on the class of all* a.c. *functions $y = y(u)$, $a \leqq u \leqq b$, having total variation $\leqq M$ and such that $(u, y(u))$ lies in A for all u.*

However, in a previous paper I have had need of a semi-continuity theorem more general than this. We therefore suppose that g satisfies the conditions

(8.1a) $g(u, y, \eta)$ is defined, finite, and l.s.c. for all (u, y) in a closed set A and all η;

(8.1b) $g(u, y, 0)$ is bounded above.

From $g(u, y, \eta)$ we form a new function f as in §4:

(8.2) $f(u, y, \xi, \eta) \equiv \xi g(u, y, \eta/\xi)$, $(u, y) \epsilon A$, $\xi > 0$, $f(u, y, 0, 0) = 0$, $(u, y) \epsilon A$,

and we again introduce the notation $x^0 = u$, $x^i = y^i$, $r^0 = \xi$, $r^i = \eta^i$. We readily see that $f(x, tr) = tf(x, r)$ if $x \epsilon A$, and $t \geqq 0$. If m is any constant, the semi-continuity of $\int f(x, \dot{x})dt$ is equivalent to that of $\int (f(x, \dot{x}) + m\dot{x}^0)dt$, since the two integrals differ only by the functional $\int m\dot{x}^0 dt$, which, being m times the difference of the final and initial values of x^0, is a continuous functional. Hence instead of (8.1b) we may assume without loss of generality that

(8.3) $f(x^0, \cdots, x^q, 1, 0, \cdots, 0) < 0$.

Let us now assume that

(8.4) for fixed $(u, y) \epsilon A$, the function $g(u, y, \eta)$ is convex in η. This implies

(8.5) for fixed $x \epsilon A$, the function $f(x, r)$ is convex on the set of all r with $r^0 > 0$.

We must show that for any such r_1 and r_2 we have

$$f(x, r_1) + f(x, r_2) \geqq 2f(x, \tfrac{1}{2}(r_1 + r_2));$$

that is,

$$\xi_1 g(u, y, \eta_1/\xi_1) + \xi_2 g(u, y, \eta_2/\xi_2) \geqq (\xi_1 + \xi_2)g(u, y, (\eta_1 + \eta_2)/(\xi_1 + \xi_2)).$$

Writing $k_i = \xi_i/(\xi_1 + \xi_2)$ $(i = 1, 2)$, we have $k_i > 0$, $k_1 + k_2 = 1$, and

$$k_1\eta_1/\xi_1 + k_2\eta_2/\xi_2 = (\eta_1 + \eta_2)/(\xi_1 + \xi_2).$$

The inequality to be proved is then

$$k_1 g(u, y, \eta_1/\xi_1) + k_2 g(u, y, \eta_2/\xi_2) \geqq g(u, y, k_1\eta_1/\xi_1 + k_2\eta_2/\xi_2),$$

in which form it is easily seen to follow from (8.4).

We next prove that if (8.3) and (8.5) hold, $f(u, y, \xi, \eta)$ is a monotonic decreasing function of ξ for fixed (u, y, η). Suppose $0 < \xi < \xi + h$. Remembering the homogeneity and continuity of f, we have

$$\begin{aligned} f(u, y, \xi + h, \eta) &= 2f(u, y, (\xi + h)/2, \eta/2) \leqq f(u, y, \xi, \eta) + f(u, y, h, 0) \\ &= f(u, y, \xi, \eta) + hf(u, y, 1, 0) < f(u, y, \xi, \eta), \end{aligned}$$

which was to be proved. Hence $f(u, y, \xi, \eta)$ tends to a limit, finite or infinite, as $\xi \to 0$, and we can define

$$f(u, y, 0, \eta) = \lim_{\xi \to 0} f(u, y, \xi, \eta). \tag{8.6}$$

This is easily seen to be consistent with the definition $f(u, y, 0, 0) = 0$.

Next we prove that $f(u, y, \xi, \eta)$ is l.s.c. For $\xi > 0$ this is an immediate consequence of (8.1a), for if $(u_n, y_n, \xi_n, \eta_n) \to (u, y, \xi, \eta)$, then

$$\begin{aligned} \liminf f(u_n, y_n, \xi_n, \eta_n) = \liminf \xi_n g(u_n, y_n, \xi_n/\eta_n) &\geqq \xi g(u, y, \eta/\xi) \\ &= f(u, y, \xi, \eta). \end{aligned}$$

Hence for every positive integer p the function $f(u, y, \xi + 1/p, \eta)$ is l.s.c. for $(u, y) \epsilon A$ and $\xi \geqq 0$. Now let $p \to \infty$. For $\xi > 0$ the function f is convex, hence continuous, so $f(u, y, \xi + 1/p, \eta) \to f(u, y, \xi, \eta)$. For $\xi = 0$ the same

relation holds by definition (8.6). As p increases and $1/p$ decreases, $f(u, y, \xi + 1/p, \eta)$ increases, since it is monotonic decreasing in ξ. Hence $f(u, y, \xi, \eta)$ is the limit of an increasing sequence of l.s.c. functions, so it is itself l.s.c. Moreover, $f(u, y, \xi + 1/p, \eta)$ is convex in (ξ, η) for $\xi \geqq 0$, so for fixed (x, y) we see that f is the limit of an increasing sequence of convex functions, so it is itself convex in (ξ, η).

Thus we have shown that if R is the set of (x, r) with $x \epsilon A$ and $r^0 \geqq 0$, conditions (2.2a) and (2.2b) are satisfied by f. Condition (2.2d) is satisfied because $f(x, r)$ is convex in r for fixed x. Condition (2.2c) holds, because, first, R is closed, and second, if $(x, y, \xi, \eta) \epsilon R$, so is $(x, y, \xi + h, \eta)$ for all $h > 0$. From Theorem 3.1 we therefore conclude:

THEOREM 8.2. *If $g(u, y, \eta)$ satisfies conditions* (8.1) *and* (8.4), *then for every $M > 0$ the integral $\int f(x, \dot{x})dt$ is l.s.c. on the class of all curves lying in A, having lengths $\leqq M$, and having $\dot{x}^0 \geqq 0$; here f is defined by* (8.2) *and* (8.6).

If in particular we restrict our attention to curves having a.c. representations of the form $y = y(u)$, $a \leqq u \leqq b$, we obtain as in §5, the following generalization of Theorem 8.1:

THEOREM 8.3. *If $g(u, y, \eta)$ satisfies conditions* (8.1) *and* (8.4), *then for every $M > 0$ the integral $\mathcal{G}(y) = \int g(u, y, \dot{y})du$ is l.s.c. on the class of all* a.c. *functions $y(u)$, $a \leqq u \leqq b$, such that $(u, y(u))$ lies in A and $y(u)$ is of total variation $\leqq M$.*

§9. If in addition to hypothesis (2.1) we assume that the set R consists of all (x, r) with x in a closed set A, and if moreover there is an $m > 0$ such that $f(x, r) \geqq m \mid r \mid$, it follows readily that in the class K of all curves joining two fixed points x_1 and x_2 there is a curve C for which $\mathcal{F}(C)$ is least. In case the lower bound i of $\mathcal{F}(C)$ on the class K is ∞, any curve of K will serve. Otherwise $0 \leqq i < \infty$. We choose a sequence $\{C_n\}$ of curves of K such that $i + 1 \geqq \mathcal{F}(C_1) \geqq \mathcal{F}(C_2) \cdots \to i$. For all of these curves we have $\mathcal{L}(C_n) = \int \mid \dot{x}_n \mid dt \leqq m^{-1} \int f(x_n, \dot{x}_n)dt = m^{-1}\mathcal{F}(C_n) \leqq (i + 1)/m$. Hence by Hilbert's theorem there is a curve of accumulation C_0, which also joins x_1 to x_2. From the C_n we choose a subsequence $\{C_\alpha\}$ with limit C_0. Since $C_0 \epsilon K$, we see that $\mathcal{F}(C_0) \geqq i$. On the other hand, by Theorem 3.1

$$i = \liminf \mathcal{F}(C_n) \geqq \mathcal{F}(C_0) \geqq i,$$

so $\mathcal{F}(C_0) = i$, and C_0 is the curve sought.

We now apply this very simple existence theorem to three special cases. As a first example, we consider a number of pieces of glass with index of refraction $\rho > 1$ suspended in a vacuum and consider the path that a light-ray would traverse in going from x_1 to x_2. The reciprocal of the velocity of light at any point is $\rho(x)/c$, where c is the velocity of light in vacuo and $\rho(x) = \rho$ if x is in the glass and $\rho(x) = 1$ if x is in the vacuum. The time of traversal of any given path $x = x(t)$ is then $c^{-1} \int \rho(x) \mid \dot{x} \mid dt$. Here $\rho(x) \mid \dot{x} \mid \geqq \mid \dot{x} \mid$, and if the glass be regarded as forming an open set $\rho(x) \mid \dot{x} \mid$ is l.s.c. The others of conditions (2.1) clearly hold. Therefore there exists a path for which the time of traversal is least, and by Fermat's principle, this is the path sought.

The theory applies equally well if we replace the glass by anisotropic crystals and require that the path have a point in common with a given point set (mirror).

As a second example we consider the Zermelo navigation problem.[15] A ship whose velocity relative to the water is k is to travel from a point x_1 to a point x_2, the water being in motion and having at the point x a velocity $(v^1(x), v^2(x))$. We suppose that the "sea" is a closed set A and that $v(x)$ is continuous. (If v were a function of both x and the time τ, this would be the general form of the navigation problem, which is a Mayer problem.) For this problem the time of traversal is $\int f(x, \dot{x})dt$, where $f(x, r)$ is defined as follows. For fixed x, a half line from the origin, $r = 0$, may meet the circle $| r - v(x) | = k$ in 0, 1, or 2 points distinct from $r = 0$. In the first case we leave $f(x, r)$ undefined on the ray. In the second case, if the intersection is at r_1, we set $f(x, tr_1) = t$ for $t \geqq 0$. In the third case, if r_1 is the intersection further from the origin we set $f(x, tr_1) = t$ for $t \geqq 0$. Thus for each x the set $E_{x,u}$ on which $f(x, r)$ is defined and $f(x, r) \leqq u$ consists of all the points of the circumference $| r - uv(x) | = uk$ and all points on the line segments joining this circumference to $r = 0$. Since $| r_1 | \leqq k + | v(x) |$, we find that whenever $f(x, r)$ is defined the inequality

$$f(x, r) \geqq | r | /(k + | v(x) |)$$

holds. The sets for which $f(x, r)$ is defined will be called "attainable."

We now define an auxiliary function $F(x, r)$ which is equal to $f(x, r)$ when (x, r) is attainable, and is ∞ if $x \in A$ and (x, r) is not attainable. This function is l.s.c. For if $u < F(x, r)$, then r is not in the set $E_{x,u}$, so it has a distance $2\delta > 0$ from that set. If $\bar{x}$ remains in a neighborhood N_x of x, the center $uv(x)$ of the circle $| r - uv(x) | = uk$ moves less than δ, so for such $\bar{x}$ the distance from r to $E_{\bar{x},u}$ is still less than δ. Thus for $\bar{x} \in N_x$ and $| \bar{r} - r | < \delta$ we have either $f(\bar{x}, \bar{r})$ defined and $F(\bar{x}, \bar{r}) = f(\bar{x}, \bar{r}) > u$, or else $f(\bar{x}, \bar{r})$ undefined and $F(\bar{x}, \bar{r}) = \infty > u$. Conditions (2.2b, c) clearly hold for F. To verify (2.2d), we first dispose of $r_0 = 0$. For this, if $u < F(x_0, r_0) = 0$, we choose $a_\alpha r^\alpha \equiv 0$, and then $F(x_0, r) \geqq a_\alpha r^\alpha$ for all r. If $r_0 \neq 0$, choose any $u < F(x_0, r_0)$. Since $F(x_0, r_0) > 0$, we can determine a positive w so that $u < w < F(x_0, r_0)$. Consider the set $E_{x_0,w}$. This is convex, and r_0 is not in it, so we can find a line not passing through the origin and separating r_0 from $E_{x_0,w}$. Let the equation of this line be written in the form $a_\alpha r^\alpha - w = 0$. The line separates r_0 from the origin, so $a_\alpha r_0^\alpha - w > 0$. The set $E_{x_0,w}$ lies on the same side of the line as the origin, so for $r \in E_{x_0,w}$ we have $a_\alpha r^\alpha - w \leqq 0$. That is, if $F(x_0, r) = w$, then $a_\alpha r^\alpha \leqq w = F(x_0, r)$. By homogeneity, the inequality $a_\alpha r^\alpha \leqq F(x_0, r)$ holds for all r, and (2.2d) is satisfied.

Since F satisfies all the hypotheses of our existence theorem, there is a curve C joining x_1 to x_2 for which $\int F(x, \dot{x})dt$ is least. If this last integral is finite, then $F(x, \dot{x})$ must be finite for almost all t. But $F(x, \dot{x}) = \infty$ unless $(x, \dot{x})$ is attainable, so for almost all t the set $(x, \dot{x})$ is attainable and $F(x, \dot{x}) = f(x, \dot{x})$. Thus we have shown that if it is possible to travel from x_1 to x_2 in a finite time,

[15] Carathéodory, *Variationsrechnung*, p. 234.

it is possible to make the journey along a path $x = x_0(t)$ such that $(x_0, \dot{x}_0)$ is attainable for almost all t and such that the time of the voyage along the path $x = x_0(t)$ is the least possible.

For our final example we consider the problem of minimizing an integral $\int f(x, \dot{x})dt$ in the class (assumed non-vacuous) of curves joining two fixed points x_1, x_2 and satisfying for almost all t a set of equations $c^i_\alpha(x)\dot{x}^\alpha = 0$, $(i = 1, \cdots, m < q)$. We assume that $f(x, r)$ is defined and continuous for all x in a closed set A and all r, and that $f(x, tr) = tf(x, r)$ for $t \geqq 0$. We further assume that the $c^i_j(x)$ are continuous on A, and that there is a number $p > 0$ such that $f(x, r) \geqq p \mid r \mid$ whenever the equations $c^i_\alpha(x)r^\alpha = 0$ are satisfied.[16] (A set (x, r) such that $c^i_\alpha(x)r^\alpha = 0$ will be called admissible.)

As before, we define $F(x, r)$ to be equal to $f(x, r)$ whenever $c^i_\alpha(x)r^\alpha = 0$, and $F(x, r) = \infty$ elsewhere. This function is lower semi-continuous. For the admissible arguments (x, r) form a closed set, so if (x_0, r_0) is not admissible, there is a neighborhood of (x_0, r_0) on which $F(x, r) = \infty$. If (x_0, r_0) is admissible, for every $\epsilon > 0$ there is a neighborhood of (x_0, r_0) on which $F(x, r) \geqq f(x, r) \geqq f(x_0, r_0) - \epsilon = F(x_0, r_0) - \epsilon$. Conditions (2.2b, c) clearly hold. With regard to (2.2d) we now assume that $\mathcal{E}_f(x, r, \bar{r}) \geqq 0$ whenever (x, r) and $(x, \bar{r})$ are admissible. Then if (x_0, r_0) is admissible, the function $r^\alpha f_{(\alpha)}(x_0, r_0)$ serves for the $a_\alpha r^\alpha$ of (2.2d); for if (x_0, r) is admissible we have $0 \leqq \mathcal{E}_f(x_0, r_0, r) \doteq f(x_0, r) - f_{(\alpha)}(x_0, r_0)r^\alpha = F(x_0, r) - a_\alpha r^\alpha$; otherwise $F(x_0, r) = \infty > a_\alpha r^\alpha$. If (x_0, r_0) is not admissible, then the equations $c^i_\alpha r^\alpha_0 = 0$ are not all satisfied; say $c^h_\alpha(x_0)r^\alpha_0 > 0$. Choose now any linear function $b_\alpha r^\alpha$ such that $F(x_0, r) \geqq b_\alpha r^\alpha$ for all admissible r; we have just seen that this is possible. Now if u be any number, it is possible to find an N large enough so that $u < b_\alpha r^\alpha_0 + Nc^h_\alpha(x_0)r^\alpha_0$. Then the function $a_\alpha r^\alpha \equiv b_\alpha r^\alpha + Nc^h_\alpha r^\alpha$ is the one sought; for if (x_0, r) is admissible, we have $F(x_0, r) = f(x_0, r) \geqq b_\alpha r^\alpha = a_\alpha r^\alpha$, while if (x_0, r) is not admissible, $F(x_0, r) = \infty > a_\alpha r^\alpha$.

Since $F(x,r)$ satisfies all the hypotheses of our existence theorem, there is a curve $C: x = x(t)$ for which $\int F(x, \dot{x})dt$ is least. But this least value is by hypothesis finite, so $F(x, \dot{x}) < \infty$ for almost all t. Therefore for almost all t the equations $c^i_\alpha(x)\dot{x}^\alpha(t) = 0$ must be satisfied, and we have proved

THEOREM 9.1. *If the functions $c^i_j(x)$ $(j = 1, \cdots, q, i = 1, \cdots, m < q)$ are continuous on a closed set A, and $f(x, r)$ is defined and continuous for $x \in A$ and all r, and $f(x, tr) = tf(x, r)$ for $t \geqq 0$, and there is a positive number p such that $f(x, r) \geqq p \mid r \mid$ whenever $c^i_\alpha(x)r^\alpha = 0$, and $\mathcal{E}(x, r, \bar{r}) \geqq 0$ whenever $x \in A$ and $c^i_\alpha(x)r^\alpha = c^i_\alpha(x)\bar{r}^\alpha = 0$, then in the class of all curves $x = x(t)$ joining two fixed points x_1, x_2 and such that $c^i_\alpha(x)\dot{x}^\alpha = 0$ for almost all t, there exists a curve $x = x_0(t)$ for which the integral $\int f(x, \dot{x})dt$ assumes its least value.*

UNIVERSITY OF VIRGINIA.

[16] Of course this last could be replaced by weaker hypotheses.

NOTE ON A SINGULAR INTEGRAL. II

By E. P. Northrop

1. Introduction. This paper is concerned with the convergence in the mean to $f(x)$, as $m \to \infty$, of the integral

$$T_m(x;f) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} K(x-u;m)\, f(u)du,$$

and is a generalization of results obtained in an earlier note by the author.[1] As a point of departure for the present note we shall, after a few preliminary remarks, introduce the main theorem (hereafter referred to as Theorem I) of the first one.

Since all of the functions to be considered will be defined over the infinite range, we shall denote the Lebesgue class $L_r(-\infty, +\infty)$ by simply L_r. We write $\| f(x) \|_r$ for the norm of a function in L_r, and define it by means of the relation

$$\| f(x) \|_r \equiv \left[\int_{-\infty}^{+\infty} | f(x) |^r dx \right]^{\frac{1}{r}}.$$

The Fourier transform of a function $f(x) \epsilon L_r$, $r > 1$, is defined (provided it exists) as the limit in the mean of order s, $1/r + 1/s = 1$, as $A \to \infty$, of the integral

$$(2\pi)^{-\frac{1}{2}} \int_{-A}^{A} e^{-ixt} f(t)dt,$$

and will be denoted by $T[x; f(t)]$, or, if there can be no confusion regarding the argument, more simply by $T[f(x)]$. The inverse Fourier transform of $f(x)$, denoted by $T^{-1}[x; f(t)] \equiv T^{-1}[f(x)]$, is defined by the same expression, except that e^{-ixt} is replaced by its complex conjugate.

Theorem I. *Let $K(x; m) \epsilon L_2$ for every m. Then in order that $T_m(x;f) \epsilon L_2$ for every m and $\| T_m(x;f) - f(x) \|_2 \to 0$ as $m \to \infty$, for every $f(x) \epsilon L_2$, it is necessary and sufficient that $K(x; m)$ satisfy the conditions*

(i) $$\underset{-\infty < x < +\infty}{\text{e.l.u.b.}} | T[K(x;m)] | = M_m, \quad \textit{and} \quad \overline{\lim_{m\to\infty}} M_m < M,$$

(ii) $$\lim_{m\to\infty} \int_a^b | T[K(x;m)] - 1 |^2 dx = 0$$

for every finite a and b.

Remarks. In condition (i), as throughout the rest of the paper, M_m is a

Received June 12, 1935.

[1] Bull. Amer. Math. Soc., vol. 40 (1934), pp. 494–496.

finite function of m, M is a constant, and e.l.u.b. denotes essential least upper bound, i.e., the least upper bound, for a fixed m, except for a set of measure zero. The reader who refers to the first note will find that condition (i) has been revised. It was originally stated thus: $|T[K(x; m)]| < M$ for all m and almost all x. In this case, Theorem I would not be true, for the purposes of necessity, when m is a continuous parameter. This difficulty was not mentioned explicitly by Lebesgue in connection with a theorem[2] upon which the proof of Theorem I is based, and the difficulty was again overlooked by the present author. In Theorem I as stated above, however, it is immaterial whether m is a continuous or a discrete parameter. The same may be said for the remainder of the theorems in this note, with the exception of Theorem IIa.

It is natural to inquire whether or not this theorem can be generalized so as to cover the case where $f(x) \in L_r$, $1 \leqq r \leqq \infty$. This note is an endeavor to answer this question, and does so partially, in that sufficient conditions for convergence are obtained in the case $2 < r < \infty$, and certain necessary conditions in the case $1 < r < 2$. The difficulties involved in the cases $1 < r < 2$ relative to sufficient conditions, and $2 < r < \infty$ relative to necessary conditions, will be discussed later. No attempt has been made here to treat the extreme cases $r = 1$, $r = \infty$. It might be pointed out in this connection that H. Hahn has obtained necessary and sufficient conditions for the case $r = 1$, although these are not in terms of the Fourier transform of the kernel.[3]

Throughout the paper it will be assumed that p and q are numbers satisfying the relations $1 < p < 2$, $1/p + 1/q = 1$. It follows that $2 < q < \infty$. We shall have occasion to use Hölder's inequality: if $f_1(x) \in L_p$, and $f_2(x) \in L_q$, then

$$\left| \int_{-\infty}^{+\infty} f_1(x) f_2(x) dx \right| \leqq \| f_1(x) \|_p \, \| f_2(x) \|_q;$$

as well as the following known properties of the Fourier transform:[4]

(a) If $f(x) \in L_p$, then $T[f(x)]$ and $T^{-1}[f(x)]$ exist and belong to L_q, and $T^{-1}\{T[f(x)]\} = T\{T^{-1}[f(x)]\} = f(x)$ almost everywhere.

(b) For every $f_1(x)$ and $f_2(x)$ belonging to L_p,

$$\int_{-\infty}^{+\infty} f_1(x) T[f_2(x)] dx = \int_{-\infty}^{+\infty} f_2(x) T[f_1(x)] dx. \tag{1.1}$$

(c) If $f(x) \in L_p$, then

$$\| T[f(x)] \|_q \leqq A(p) \, \| f(x) \|_p, \tag{1.2}$$

where $A(p)$ is a finite quantity depending only upon p.

[2] Ann. de la Fac. des Sc. de l'Univ. de Toulouse, (3), vol. 1 (1909), p. 52.

[3] Kais. Ak. der Wiss. in Wien, Denkschriften, vol. 93 (1917), p. 667.

[4] For property (a) see E. Hille and J. D. Tamarkin, Bull. Amer. Math. Soc., vol. 39 (1933), pp. 768–774; for properties (b) and (c), E. C. Titchmarsh, Proc. Lond. Math. Soc., (2), vol. 23 (1924–25), pp. 288 and 287 resp.

2. Sufficient conditions in the case $f(x) \epsilon L_q$. We have two main theorems in this case; in the first, $f(x)$ belongs to a (dense) subset of L_q, and in the second, $f(x)$ is an arbitrary function of L_q.

THEOREM II. *In order that* $T_m(x; f) \epsilon L_q$ *for every* m *and*

$$\| T_m(x; f) - f(x) \|_q \to 0 \text{ as } m \to \infty,$$

for every $f(x)$ *which is the Fourier transform in* L_q *of some function in* L_p, *it is sufficient that* $K(x; m)$ *satisfy the conditions*

(i) $$K(x; m) \epsilon L_p \text{ for every } m,$$

(ii) $$\underset{-\infty<x<+\infty}{\text{e.l.u.b.}} \mid T[K(x; m)] \mid = M_m, \quad \text{and} \quad \overline{\lim_{m\to\infty}} M_m < M,$$

(iii) $$\lim_{m\to\infty} \int_a^b \mid T[K(x; m)] - 1 \mid^p dx = 0$$

for every finite a *and* b.

Proof. Conditions (ii) and (iii) are obviously equivalent to those obtained by replacing T by T^{-1}. In the following we shall use the conditions so revised. We show first that $T_m(x; f) \epsilon L_q$. If in (1.1) we put $f_1(u) = K(x - u; m)$, $T[f_2(u)] = f(u)$, then

$$\begin{aligned} T_m(x; f) &= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} T^{-1}[f(u)]\, T[u; K(x - t; m)]du \\ &= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{-iux}\, T^{-1}[f(u)]\, T^{-1}[K(u; m)]du, \end{aligned}$$

since it can be easily verified that $T[u; K(x - t; m)] = e^{-iux}T^{-1}[u; K(t; m)] \equiv e^{-iux}T^{-1}[K(u; m)]$. Then by condition (ii), $\mid T^{-1}[f(u)]T^{-1}[K(u; m)] \mid < M \mid T^{-1}[f(u)] \mid$. As the right term of this inequality belongs to L_p, so also must the left term, and $T_m(x; f)$ can be regarded as the Fourier transform in L_q of $T^{-1}[f(x)]T^{-1}[K(x; m)]$.

Keeping in mind that $f(x)$ can be thought of as the Fourier transform of $T^{-1}[f(x)]$, we can then write, with the aid of (1.2),

(2.1) $$\| T_m(x; f) - f(x) \|_q \leqq A(p) \| T^{-1}[f(x)]\{T^{-1}[K(x; m)] - 1\} \|_p.$$

The norm on the right of this last relation is the p-th root of the integral

$$\int_{-\infty}^{+\infty} \mid T^{-1}[f(x)] \mid^p \mid T^{-1}[K(x; m)] - 1 \mid^p dx.$$

By the theorem of Lebesgue used in the proof of Theorem I, the conditions (ii) and (iii) are sufficient for the convergence of the above integral to zero, as $m \to \infty$. This, by (2.1), proves the theorem.

THEOREM III. *In order that* $T_m(x; f) \epsilon L_q$ *for every* m *and*

$$\| T_m(x; f) - f(x) \|_q \to 0$$

as $m \to \infty$, *for every* $f(x) \epsilon L_q$, *it is sufficient that* $K(x; m)$ *satisfy conditions* (i) *and* (iii) *of Theorem* II, *and the condition*

$$\text{(ii}'\text{)} \qquad \int_{-\infty}^{+\infty} |K(x; m)|\, dx < N \text{ (a constant) for every } m.$$

Proof. Note that condition (ii) of Theorem II has been strengthened, as (ii′) implies (ii). We first show, as before, that $T_m(x; f) \epsilon L_q$. To do so, we use (ii′) and Hölder's inequality in a somewhat modified form. We have, keeping in mind that $1/p + 1/q = 1$,

$$\left| (2\pi)^{\frac{1}{2}} T_m(x; f) \right| \leqq \int_{-\infty}^{+\infty} |K(x-u; m)|^{\frac{1}{p}} \left\{ f(u) |K(x-u; m)|^{\frac{1}{q}} \right\} du$$

$$\leqq \left[\int_{-\infty}^{+\infty} |K(x-u; m)|\, du \right]^{\frac{1}{p}} \left[\int_{-\infty}^{+\infty} |f(u)|^q |K(x-u; m)|\, du \right]^{\frac{1}{q}}$$

$$< N^{\frac{1}{p}} \left[\int_{-\infty}^{+\infty} |f(u)|^q |K(x-u; m)|\, du \right]^{\frac{1}{q}}.$$

That is,

$$\text{(2.2)} \qquad \| T_m(x; f) \|_q < (2\pi)^{-\frac{1}{2}} N \| f(x) \|_q .$$

We now make use of the fact that, given an arbitrary function $f(x) \epsilon L_q$, and an $\epsilon > 0$, we can find a function $\varphi(x) \epsilon L_q$ which is the Fourier transform of a function in L_p and such that

$$\text{(2.3)} \qquad \| f(x) - \varphi(x) \|_q < \epsilon.$$

(It would be sufficient to use, as $\varphi(x)$, a step-function; for the Fourier transform of such a function is in every L_r, $r > 1$.) With this in mind, we write

$$\| T_m(x; f) - f(x) \|_q \leqq \| \varphi(x) - f(x) \|_q + \| T_m(x; f) - T_m(x; \varphi) \|_q + \| T_m(x; \varphi) - \varphi(x) \|_q .$$

If we can show that the right side of this inequality can be made arbitrarily small by the choice of a sufficiently large m, then the left side will also have this property. The first term is arbitrarily small by (2.3), as is the second term. For in view of the fact that T_m is an additive transformation,

$$\| T_m(x; f) - T_m(x; \varphi) \|_q = \| T_m(x; f - \varphi) \|_q < (2\pi)^{-\frac{1}{2}} N \| f - \varphi \|_q,$$

by (2.2). Finally, to the third term we can, for a fixed φ, apply Theorem II. This proves the theorem.

3. Necessary conditions in the case $f(x) \epsilon L_p$.

THEOREM IV.[5] *For every* m *let* $K(x; m)$ *be the Fourier transform in* L_q *of some function in* L_p. *Then in order that* $T_m(x; f) \epsilon L_p$ *for every* m *and*

[5] The author wishes to express his indebtedness to the referee for suggesting this theorem, and to E. Hille for indicating the argument which replaces the theorem of Lebesgue.

$\| T_m(x; f) - f(x) \|_p \to 0$ *as* $m \to \infty$, *for every* $f(x) \in L_p$, *it is necessary that* $K(x; m)$ *satisfy the conditions*

(i) $$\underset{-\infty < x < +\infty}{\text{e.l.u.b.}} \, | T[K(x; m)] | = M_m, \quad \textit{and} \quad \overline{\lim_{m \to \infty}} M_m < M,$$

(ii) $$\lim_{m \to \infty} \int_a^b | T[K(x; m)] - 1 |^q \, dx = 0$$

for every finite a and b.

Proof. If in (1.1) we put $f_1(u) = f(u)$, $T[f_2(u)] = K(x - u; m)$, then

$$T_m(x; f) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} T[f(u)] \, T^{-1}[u; K(x - t; m)] \, du$$
$$= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{iux} T[f(u)] \, T[K(u; m)] \, du,$$

since it can be easily verified that $T^{-1}[u; K(x - t; m)] = e^{iux} T[u; K(t; m)] \equiv e^{iux} T[K(u; m)]$. From the above it is evident that $T_m(x; f)$ can be regarded as the inverse Fourier transform of $T[f(x)] \, T[K(x; m)]$. If now we assume that $T_m(x; f)$ belongs to L_p, it follows that its Fourier transform in L_q is almost everywhere equal to $T[f(x)] T[K(x; m)]$. We can then write, with the aid of (1.2),

(3.1) $$\| T_m(x; f) \|_p \geqq [A(p)]^{-1} \| T[f(x)] \, T[K(x; m)] \|_q,$$

(3.2) $$\| T_m(x; f) - f(x) \|_p \geqq [A(p)]^{-1} \| T[f(x)] \{ T[K(x; m)] - 1 \} \|_q.$$

If the term on the left of (3.2) tends to zero as $m \to \infty$, so also must the term on the right. Similarly, the boundedness of the left side of (3.1) implies that of the right side. We cannot apply the theorem of Lebesgue directly to the present situation, as we did in Theorem II, because $T[f(x)]$ is not an arbitrary function of L_q. This follows from the fact that the Fourier transform maps L_p on a (dense) subset of L_q. This difficulty, however, can be taken care of as follows, by an argument of Lebesgue's type.

For any choice of a and b, $a < b$, $b - a < \infty$, the function

$$f(x, a, b) = \frac{e^{ibx} - e^{iax}}{ix \sqrt{2\pi}}$$

is in L_p, and

$$\| f(x, a, b) \|_p = \alpha(p)(b - a)^{\frac{1}{q}},$$

where $\alpha(p)$ depends only upon p. Furthermore,

$$T[f(x, a, b)] = \begin{cases} 1, \; a < x < b, \\ \frac{1}{2}, \; x = a, b, \\ 0, \text{ elsewhere.} \end{cases}$$

For this function, (3.2) becomes

$$\| T_m(x;f) - f(x) \|_p \geqq [A(p)]^{-1}\left[\int_a^b | T[K(x;m)] - 1 |^q\, dx\right]^{\frac{1}{q}}.$$

Hence condition (ii) is necessary.

Similarly, for the function $f(x, a, b)$, (3.1) gives

$$\| T_m(x;f) \|_p \geqq [A(p)\alpha(p)]^{-1}\left\{\frac{1}{b-a}\int_a^b | T[K(x;m)] |^q\, dx\right\}^{\frac{1}{q}} \| f(x, a, b) \|_p.$$

As we shall presently prove, the transformation $T_m(x; f)$ is bounded in L_p for every m; i.e., there exists a finite $B(m)$ such that $\| T_m(x;f) \|_p \leqq B(m) \| f(x) \|_p$ for every $f(x) \epsilon L_p$. This implies that

$$\frac{1}{b-a}\int_a^b | T[K(x;m)] |^q\, dx \leqq [M(m)]^q, \tag{3.3}$$

where $M(m)$ is a finite quantity independent of a and b. But if a is fixed, and $b \to a$, then, by a well-known theorem of Lebesgue, the left side of (3.3) tends to $| T[a; K(t; m)] |^q$ for almost all values of a. Hence

$$| T[K(x; m)] | \leqq M(m) \tag{3.4}$$

for almost all x.

In order to prove the boundedness of $T_m(x; f)$ in L_p, we note first that Hölder's inequality gives

$$| T_m(x;f) | \leqq (2\pi)^{-\frac{1}{2}} \| K(x; m) \|_q \| f(x) \|_p, \tag{3.5}$$

so that the transformation is bounded for every m. Now define the function

$$T_m^n(x;f) = \begin{cases} T_m(x;f) & | x | \leqq n, \\ 0, & \text{elsewhere.} \end{cases}$$

By assumption, $T_m(x; f) \epsilon L_p$, so that à fortiori $T_m^n(x; f) \epsilon L_p$. But $T_m^n(x; f)$ is a bounded transformation in L_p, since by (3.5),

$$\| T_m^n(x;f) \|_p \leqq (2\pi)^{-\frac{1}{2}}(2n)^{\frac{1}{p}} \| K(x; m) \|_q \| f(x) \|_p.$$

Furthermore, $\| T_m^n(x; f) - T_m(x; f) \|_p \to 0$ as $n \to \infty$ for every $f(x) \epsilon L_p$. But if a sequence of bounded linear transformations converges at every point of L_p, then the bounds of the transformations must be uniformly bounded, and the limiting transformation is bounded.[6] Hence $T_m(x; f)$ is bounded in L_p for every m.

The same argument shows that the bounds $M(m)$ of (3.4) must be uniformly bounded, since by assumption $T_m(x;f)$ converges all over the space L_p. Hence condition (i) is necessary. This completes the proof of the theorem.

[6] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 80, Th. 5.

4. **Special cases of Theorems II, III, and IV.** It is perhaps of some interest to see what simplifications are brought about in the conditions on the kernel in case it is of the form $K(x; m) = mk(mx)$. We have the following theorems.

THEOREM IIa. *If in Theorem* II, $K(x; m) = mk(mx)$, *conditions* (ii) *and* (iii) *can be replaced, respectively, by the following*:

$$\text{(iia)} \qquad | T[k(x)] | < M \text{ for almost all } x,$$

$$\text{(iiia)} \qquad \lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} | T[k(x)] - 1 |^p \, dx = 0,$$

where h is a continuous parameter.

Remarks. It will be shown that (ii) and (iia) are equivalent in the case under consideration, and also that (iii) and (iiia) are equivalent provided m tends continuously to ∞. If, on the other hand, m runs over an arbitrary sequence, (iiia) appears to be more stringent than (iii).

Proof. It is easily verified that

$$T[x; mk(mt)] = T[x/m; k(t)], \tag{4.1}$$

whereupon (ii) and (iia) are obviously equivalent.

As for (iii), it becomes, in view of (4.1),

$$\lim_{m \to \infty} \int_a^b | T[x/m; k(t)] - 1 |^p \, dx = 0$$

for every finite a and b. That is,

$$\lim_{m \to \infty} m \int_{\frac{a}{m}}^{\frac{b}{m}} | T[k(x)] - 1 |^p \, dx = 0;$$

or, putting $m = 1/n$ and dividing by 2,

$$\lim_{n \to 0} \frac{1}{2n} \int_{an}^{bn} | T[k(x)] - 1 |^p \, dx = 0 \tag{4.2}$$

for every finite a and b. We now show that if m (and consequently n) is a continuous parameter, (4.2) and (iiia) are equivalent. First, assume (4.2) and take $a = -1$, $b = 1$. This gives (iiia), if we substitute h for n. Next, (iiia) implies

$$\lim_{h \to 0} \frac{1}{2h} \int_{-ah}^{ah} | T[k(x)] - 1 |^p \, dx = 0; \qquad \lim_{h \to 0} \frac{1}{2h} \int_{-bh}^{bh} | T[k(x)] - 1 |^p \, dx = 0 \tag{4.3}$$

for every finite a and b. But

$$\int_{-bh}^{bh} - \int_{-ah}^{ah} = \int_{ah}^{bh} + \int_{-bh}^{-ah}.$$

Hence (4.3) implies

$$\lim_{h \to 0} \frac{1}{2h} \left\{ \int_{ah}^{bh} + \int_{-bh}^{-ah} \right\} | T[k(x)] - 1 |^p \, dx = 0.$$

Since $| T[k(x)] - 1 |^p$ is a non-negative function, each of the integrals in the last relation is non-negative or non-positive, according as $a < b$ or $a > b$, and consequently both of them must vanish in the limit. This gives (4.2), as we wished to show.

THEOREM IIIa. *If in Theorem* III, $K(x;\, m) = mk(mx)$, *conditions* (ii′) *and* (iii) *can be replaced by*

(ii′a) $$\int_{-\infty}^{+\infty} | k(x) |\, dx < N;$$

(iiia) *as in Theorem* IIa.

The proof is immediate.

THEOREM IIIb. *If in Theorem* IIIa *we desire the conditions to involve only the kernel and not its Fourier transform, we may replace* (iiia) *by the condition*

(iiib) $$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} k(x)\, dx = 1.$$

Proof. We shall show that (ii′a) and (iiib) imply (iiia). We know from (ii′a) that $k(x) \in L_1$. Hence

$$T[k(x)] = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{-ixu} k(u)\, du,$$

$$T[k(0)] = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} k(u)\, du = 1.$$

Since $T[k(x)]$ is continuous and has the value 1 at the origin, so also must its mean value. That is, the condition (iiia) must hold.

THEOREM IVa. *If in Theorem* IV, $K(x;\, m) = mk(mx)$, *conditions* (i) *and* (ii) *become respectively*

(ia) $$| T[k(x)] | < M \textit{ for almost all } x,$$

(iia) $$\lim_{n \to 0} \frac{1}{2n} \int_{-n}^{n} | T[k(x)] - 1 |^q\, dx = 0,$$

where n is defined as $1/m$.

For the proof of this theorem the reader is referred to that of Theorem IIa, where the essentials are to be found.

5. **Remarks.** Superficially it would seem that there should be theorems similar to Theorems II and III for the case $f(x) \in L_p$, and similar to IV for the case $f(x) \in L_q$. Whether or not this is true the author is not prepared to say. On the other hand, he can say with certainty that the same methods of proof would not apply, due to the asymmetrical properties of the Fourier transform, as evinced by the relation $\| T[f(x)] \|_q \leqq A(p) \| f(x) \|_p$, which was so necessary to the methods used here. An example, however, of what can be done in this direction will be given without proof. It amounts to replacing the inequality

just mentioned by a less useful and less known one due to Hardy and Littlewood;[7] to wit, if $f(x) \epsilon L_p$, then $\| T[f(x)]x^{(p-2)/p} \|_p \leqq A(p) \| f(x) \|_p$. Using this relation, the following theorem can be proved.

THEOREM A. *In order that* $\| \{T_m(x; f) - f(x)\}x^{(p-2)/p} \|_p \to 0$ *as* $m \to \infty$, *for every* $f(x)$ *which, together with its Fourier transform, belongs to* L_p, *it is sufficient that* $K(x; m)$ *satisfy the conditions*

(i) $K(x; m)$ *is for every* m *the Fourier transform in* L_q *of some function in* L_p,

(ii) *and* (iii) *as in Theorem* II.

YALE UNIVERSITY AND THE HOTCHKISS SCHOOL.

[7] Math. Annalen, vol. 97 (1926–27), p. 203.

SYMMETRIC FUNCTIONS OF NON-COMMUTATIVE ELEMENTS

By Margarete C. Wolf

Introduction. A study of symmetric polynomials of matrices, for which the commutative law of multiplication need not necessarily be valid, led to the study of symmetric polynomials of certain abstract elements for which the processes of addition and multiplication obey the postulates of a linear associative algebra. This results in a generalization of the definition of the elementary symmetric functions. For example, if x_1 and x_2 are such elements, let x_1x_2 symbolize x_1 multiplied on the right by x_2 and let $x_1 + x_2$ indicate addition of x_1 and x_2; then since x_1x_2 differs in general from x_2x_1, the second elementary symmetric function of the elements x_1 and x_2 becomes

$$E_2 = \Sigma x_1x_2 = x_1x_2 + x_2x_1;$$

but as before, $E_1 = \Sigma x_1 = x_1 + x_2$. The simple symmetric functions of third degree of the elements x_1 and x_2 are

$$\Sigma x_1x_2x_1 = x_1x_2x_1 + x_2x_1x_2, \qquad \Sigma x_1x_2^2 = x_1x_2^2 + x_2x_1^2,$$
$$\Sigma x_1^2x_2 = x_1^2x_2 + x_2^2x_1, \qquad \Sigma x_1^3 = x_1^3 + x_2^3.$$

These functions cannot be expressed as polynomials in E_1 and E_2 as in the case of commutative elements, but another polynomial, for example $\Sigma x_1x_2x_1$, must be defined as an elementary symmetric function in addition to E_1 and E_2 if the fundamental theorem is to be reëstablished. Note also that E_1E_2 differs from E_2E_1 for non-commutative elements. If three elements x_1, x_2, x_3 are considered, two polynomials of third degree are required to serve as elementary symmetric functions instead of the one function $E_3 = \Sigma x_1x_2x_3$ of the commutative elements. Two polynomials which may be used are $E_3 = \Sigma x_1x_2x_3$ and $\Sigma x_1x_2x_1$. This paper shows that as the number of elements and the degree are increased, an infinite sequence of symmetric polynomials, consisting of a finite set of one or more for each degree can be chosen so that every symmetric polynomial may be expressed uniquely in terms of the polynomials of this sequence and the coefficients of the original polynomial, with coefficients which are integral. This sequence may be chosen in more than one way but the number for each degree is unique.

Since by the Poincaré equivalence theorem every linear associative algebra is equivalent to a matric algebra, no generality is lost if the elements are taken as matrices.

1. Simple symmetric polynomials and elements completely non-commutative of order *m*. The usual definitions and theorems which apply to sym-

Received October 31, 1935; presented to the American Mathematical Society, September 10, 1935, under the title *Symmetric functions of matrices*.

metric polynomials of commutative elements must be modified in some instances in order to be applicable to symmetric polynomials of non-commutative elements. This is done in §§1, 2, and 3, and several new definitions are incorporated, giving rise to new theorems which facilitate proving in §4 the generalized fundamental theorem on symmetric functions.

A polynomial in n elements $x_1, x_2, \cdots, x_n$ is said to be symmetric in these elements if it is unaltered by any interchange of the elements. A necessary and sufficient condition that a polynomial be symmetric is that it be unchanged by every interchange of two elements.

If $x_1, x_2, \cdots, x_n$ are n non-commutative elements, a simple symmetric polynomial is defined as the sum of all terms obtained from a term $x_{i_1}^{\nu_1} x_{i_2}^{\nu_2} \cdots x_{i_k}^{\nu_k}$ by allowing the distinct subscripts to take on all possible permutations chosen from the numbers $1, 2, \cdots, n$, where $\nu_1, \nu_2, \cdots, \nu_k$ is a fixed set of exponents. Let the symbol $S_m^{(j)}(x_n) = \Sigma x_{i_1}^{\nu_1} x_{i_2}^{\nu_2} \cdots x_{i_k}^{\nu_k}$ denote a simple symmetric polynomial of the n elements $x_1, x_2, \cdots, x_n$ of degree $m = \nu_1 + \nu_2 + \cdots + \nu_k$, where j takes on one of the values $1, 2, \cdots, j_m$ for each different simple symmetric polynomial of degree m.

The sum, difference, and product of any two symmetric polynomials are symmetric. The degree of every term in the product of two simple symmetric polynomials is equal to the sum of the degrees of the two polynomials.

In general, in the total matric algebra of order n, the terms of a simple symmetric polynomial as defined above will not be distinct, because some powers and products of the matrices will be commutative and polynomial relationships will exist among the products yielding a reduction in the number of different terms.

A set of elements $x_1, x_2, \cdots, x_n$ is said to be completely non-commutative of order m if no product is commutative with any other product whenever the sum of the degrees of the two products is less than or equal to m, where the factors of each product need not be distinct. Furthermore, this set is said to be independent of order m if a polynomial of degree less than or equal to m equals another polynomial of degree less than or equal to m only if the coefficients of like terms are equal.

In this paragraph it will be demonstrated that for every $n \geqq 2$ and every m there always exists a matric algebra from which there can be chosen n matrices such that these n matrices are completely non-commutative of order m and independent of order m. Form all possible products through degree m of the n letters $x_1, x_2, \cdots, x_n$ and with unity define these as basal elements of a finite linear associative algebra A over a field K. In A let every element of degree greater than m be defined as zero. That is, let the basal elements be

$$u_0 = 1, \quad u_1 = x_1, \quad u_2 = x_2, \quad \cdots, \quad u_n = x_n, \quad u_{n+1} = x_1x_2, \quad \cdots,$$

where the multiplication table by definition is of the form $u_iu_j = u_k$ and $u_iu_j = 0$ if the degree of u_iu_j is greater than m. This finite linear associative algebra with a principal unit is equivalent to a matric algebra with basal elements determined by the above multiplication table. With the use of this theorem

one can build n matrices $x_1, x_2, \cdots, x_n$ which are independent of order m and are completely non-commutative of order m.

From the definition of n elements which are independent of order m and completely non-commutative of order m, it follows that if any symmetric polynomial of these n elements contains the term $x_{i_1}^{\nu_1} x_{i_2}^{\nu_2} \cdots x_{i_k}^{\nu_k}$, then it contains every term obtained from it by the interchange of any two elements, and therefore a term obtained by any permutation of the elements. Consequently every symmetric polynomial of degree m can be expressed as a sum of simple symmetric polynomials of degree m. In the remainder of this paper, unless specifically stated otherwise, every set of elements studied will be considered independent of order m and completely non-commutative of order m.

It is necessary and sufficient to have m such elements to express all possible simple symmetric polynomials $S_m^{(j)}(x_n)$ for each degree m, inasmuch as there are exactly m positions to fill in each term of every $S_m^{(j)}(x_n)$.

In order to calculate the number of $S_m^{(j)}(x_n)$ for every m, choose in each of these simple symmetric polynomials, $S_m^{(j)}(x_n)$, the following typical term $x_{i_1}^{\nu_1} x_{i_2}^{\nu_2} \cdots x_{i_k}^{\nu_k}$, such that in the sequence of subscripts reading from left to right, each subscript which differs from all those which precede it is the smallest integer which differs from those integers which precede it. That is, $i_{j+1} = 1, 2, \cdots, k, \cdots, j$, or $j + 1$, but $i_{j+1} = k$ only if each of the integers $1, 2, \cdots, k - 1$ have occurred at least once as a preceding subscript. Hence $i_1 = 1$; $i_2 = 2$; $i_3 = 1$ or 3; $i_4 = 1, 2, 3$, or 4, but i_4 can equal 4 only if $i_3 = 3$; $i_5 = 1, 2, 3, 4$, or 5, but i_5 can equal 5 only if $i_4 = 4$, and $i_5 \neq 4$ if, for example, $i_3 = 1$ and $i_4 = 2$. Let $n_m(x_k)$ be the number of such typical terms of degree m with k different subscripts in a term, that is, k distinct elements in a term. To calculate the total number of $S_m^{(j)}(x_n)$ for a given m, one need but compute the value of $n_m(x_k)$ for $k = 1, 2, \cdots, m$. These numbers can be obtained by means of recursion formulas. Assume that all $n_{m-1}(x_k)$, $k = 1, 2, \cdots, m - 1$, are known. A term of degree m in one element, that is x_1^m, can be obtained from a term of degree $m - 1$ in one element by taking the m-th factor x_1. A term of degree m in two distinct elements x_1, x_2 can be obtained from a term of degree $m - 1$ in one element by taking the m-th factor x_2, or from a term of degree $m - 1$ in two distinct elements, by taking the m-th factor either x_1 or x_2. Continuing in this manner, a term of degree m in k distinct elements can be obtained from a term of degree $m - 1$ in $k - 1$ distinct elements, by taking the k-th factor x_k, or from a term of degree $m - 1$ in k distinct elements by taking the k-th factor any one of the k elements $x_1, x_2, \cdots, x_k$. That is,

$$
\begin{aligned}
n_m(x_1) &= n_{m-1}(x_1) = 1, \\
n_m(x_2) &= n_{m-1}(x_1) + 2n_{m-1}(x_2), \\
n_m(x_3) &= n_{m-1}(x_2) + 3n_{m-1}(x_3), \\
&\cdot \qquad \cdot \qquad \cdot \qquad , \\
n_m(x_k) &= n_{m-1}(x_{k-1}) + kn_{m-1}(x_k), \\
&\cdot \qquad \cdot \qquad \cdot \qquad , \\
n_m(x_m) &= n_{m-1}(x_{m-1}) = 1,
\end{aligned}
$$

where the total number of polynomials $S_m^{(j)}(x_n)$ is equal to $\sum_{k=1}^{m} n_m(x_k)$. For $m = 1$ and $m = 2$, the formulas are $n_1(x_1) = 1$, $n_2(x_1) = 1$, and $n_2(x_2) = 1$.

TABLE OF THE NUMBER OF SIMPLE SYMMETRIC POLYNOMIALS FOR THE DEGREES $1, 2, \cdots, 8$

Degree

Number of distinct elements in a term	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2		1	3	7	15	31	63	127
3			1	6	25	90	301	966
4				1	10	65	350	1701
5					1	15	140	1050
6						1	21	266
7							1	28
8								1
Total	1	2	5	15	52	203	877	4140

A product of simple symmetric polynomials is equal to a sum of simple symmetric polynomials with coefficients of positive unity. To prove this statement, consider the product of two, $S_m^{(j)}(x_n) \cdot S_n^{(k)}(x_n)$, of degree m and n respectively. Every term in each of the polynomials $S_m^{(j)}(x_n)$ and $S_n^{(k)}(x_n)$ has a coefficient of positive unity by definition. If $(x_{i_1}^{\nu_1} \cdots x_{i_k}^{\nu_k})(x_{i_l}^{\nu_l} \cdots x_{i_m}^{\nu_m})$ is a term in the product, it can arise only once, since $x_{i_1}^{\nu_1} \cdots x_{i_k}^{\nu_k}$ occurs only once in $S_m^{(j)}(x_n)$, and $x_{i_l}^{\nu_l} \cdots x_{i_m}^{\nu_m}$ occurs only once in $S_m^{(k)}(x_n)$; furthermore, the term can not arise from a different factorization because the first factor must always be of degree m and that of the second, n. The argument can be extended immediately to the product of three or more polynomials.

2. **Fundamental sets of order** m. A set of simple symmetric polynomials is said to be a fundamental set of order m if every symmetric polynomial of degree m can be expressed uniquely as a polynomial in the polynomials of this set.

The existence of fundamental sets will be proved later by the construction of particular fundamental sets.

If m elements are chosen which are independent of order m and completely non-commutative of order m, then a fundamental set of order m is analogous to the elementary symmetric functions of commutative elements. Since there are a finite number of the polynomials $S_m^{(j)}(x_n)$ for each degree m, there can be only a finite number of fundamental polynomials for each degree m. In general, a fundamental polynomial of degree m in the elements $x_1, x_2, \cdots, x_n$ shall be denoted by $F_m^{(j)}$, where j identifies the different polynomials for one degree m. To each $F_m^{(j)}$ a weight m shall be assigned. Two polynomials of equal weight are said to be isobaric. The weight of a product of fundamental polynomials is equal to the sum of the weights. Since the representation of the $S_m^{(j)}(x_n)$ in terms of the fundamental polynomials is to be unique, the $F_m^{(j)}$ must be so chosen that two polynomials of the $F_m^{(j)}$ can only be equal if the coefficients of like terms are equal.

3. **An order for symmetric polynomials.** In a manner similar to that used with commutative elements, the symmetric polynomials are ordered to simplify the problem of expressing the polynomials $S_m^{(j)}(x_n)$ as polynomials in the $F_m^{(j)}$.

All simple symmetric polynomials of degree m are said to be of higher order than all those of degree less than m.

The terms of a simple symmetric polynomial are ordered in the following manner. The term $x_{i_1}^{\nu_1}x_{i_2}^{\nu_2}\cdots x_{i_k}^{\nu_k}$ is said to be of higher order than $x_{i'_1}^{\nu_1}x_{i'_2}^{\nu_2}\cdots x_{i'_k}^{\nu_k}$ if the first non-zero difference in subscripts $i - i'$ is less than zero.

The simple symmetric polynomials of degree m are ordered in the following manner. If $x_{i_1}^{\nu_1}x_{i_2}^{\nu_2}\cdots x_{i_k}^{\nu_k}$ is the highest ordered term of $S_m^{(i)}(x_n)$ and $x_{i'_1}^{\nu'_1}x_{i'_2}^{\nu'_2}\cdots x_{i'_k}^{\nu'_k}$ is the highest ordered term of $S_m^{(j)}(x_n)$, then $S_m^{(i)}(x_n)$ is said to be of higher order than $S_m^{(j)}(x_n)$ if the first non-zero difference in exponents $\nu - \nu'$ is greater than zero. In case all differences are zero, that is, all exponents are equal, $S_m^{(i)}(x_n)$ is of higher order than $S_m^{(j)}(x_n)$ if the first non-zero difference in subscripts $i - i'$ is less than zero.

A symmetric polynomial P_1 is said to be of higher order than a symmetric polynomial P_2 if, after those simple symmetric polynomials which occur in both have been deleted, the highest ordered remaining simple symmetric polynomial of P_1 is of higher order than the highest ordered remaining simple symmetric polynomial of P_2.

From this method of ordering all polynomials and their terms, a theorem, essential to later development, may be deduced, namely:

The highest ordered term of the product $S_m^{(i)}(x_k)S_n^{(j)}(x_k)$ is the highest ordered term of $S_m^{(i)}(x_k)$ multiplied on the right by that highest ordered term in $S_n^{(j)}(x_k)$ of which the first letter is the same as the last letter of the term of $S_m^{(i)}(x_k)$.

4. **A fundamental set of order m.** In the following paragraph a particular fundamental set of order m of simple symmetric polynomials will be defined. To distinguish these from other fundamental sets, designate the polynomials as $E_m^{(j)}$, where m indicates the degree and j takes on one of the values $1, 2, \cdots, j_m$

for every distinct fundamental polynomial of a fixed degree m. It is proved that the elementary symmetric functions of commutative elements are a special case of the polynomials $E_m^{(j)}$, and that the fundamental theorem concerning elementary symmetric functions may be restated and proved for the polynomials $E_m^{(j)}$.

(a) Define $E_1 = \Sigma x_1$.

(b) Every simple symmetric polynomial of degree m whose highest ordered term is not the highest ordered term of some product of weight m of $E_n^{(j)}$, where $n = 1, 2, \cdots, m-1$ and $j = 1, 2, \cdots, j_n$, is to be defined as an $E_m^{(j)}$, $j = 1, 2, \cdots, j_m$.

All $E_m^{(j)}$ through the fourth degree are given in the following set:

$$E_1 = \Sigma x_1,$$

$$E_2 = \Sigma x_1 x_2,$$

$$E_3^{(1)} = \Sigma x_1 x_2 x_1, \qquad E_3^{(2)} = \Sigma x_1 x_2 x_3,$$

$$E_4^{(1)} = \Sigma x_1 x_2^2 x_3, \qquad E_4^{(2)} = \Sigma x_1 x_2 x_1 x_2, \qquad E_4^{(3)} = \Sigma x_1 x_2 x_1 x_3,$$

$$E_4^{(4)} = \Sigma x_1 x_2 x_3 x_1, \qquad E_4^{(5)} = \Sigma x_1 x_2 x_3 x_2, \qquad E_4^{(6)} = \Sigma x_1 x_2 x_3 x_4.$$

TABLE FOR THE NUMBER OF $E_m^{(j)}$ FOR $m = 1, 2, \cdots, 6$

Degree

Number of distinct elements in a term	1	2	3	4	5	6
1	1					
2		1	1	1	1	1
3			1	4	12	33
4				1	8	44
5					1	13
6						1
Total	1	1	2	6	22	92

In constructing the $E_m^{(j)}$ for a given degree m, the following facts are useful. The first is that every simple symmetric polynomial of the form $\Sigma x_{i_1} x_{i_2} \cdots x_{i_m}$ such that $i_j \neq i_{j+1}$ is among the $E_m^{(j)}$. These polynomials are the direct generalization of the elementary symmetric functions but they are not sufficient to form a fundamental set of order m. The second useful fact is that the exponent of the first and last letter of every $E_m^{(j)}$ is unity. For suppose the

first exponent is not unity but ν_1. One can then factor the highest ordered term into $x_1^{\nu_1-1}$ and $x_1x_2^{\nu_2}\cdots x_k^{\nu_k}$. But $x_1^{\nu_1-1}$ is the highest ordered term in $(E_1)^{\nu_1-1}$. Therefore the highest ordered term of $E_m^{(j)}$ is the highest ordered term in a product $(E_1)^{\nu_1-1}$ times some other product of $E_n^{(j)}$'s. Similarly, a contradiction is reached if the last letter is assumed to have an exponent not unity.

THEOREM. *If the elements $x_1, x_2, \cdots, x_n$ are commutative, the definition of the fundamental set of order m which gives the $E_m^{(j)}$ yields the first n elementary symmetric functions.*

As in the case of commutative elements, $E_1 = \Sigma x_1$, $E_2 = \Sigma x_1x_2$. If the elements are commutative, the highest ordered term of a product is equal to the product of the highest ordered terms. Therefore the highest ordered term can be written as $x_1^{h_1}x_2^{h_2}\cdots x_k^{h_k}$, where $h_1 \geqq h_2 \geqq \cdots \geqq h_k$ and

$$h_1 + h_2 + \cdots + h_k = m.$$

If $h_1 > 1$, the highest ordered term can always be expressed as the highest ordered term of the product of elementary symmetric functions, each of weight less than m, in the following manner: $E_k^{h_k}\cdot E_{k-i}^{h_{k-i}-h_k}\cdots$.

Example. The highest ordered term of $\Sigma x_1^5x_2^5x_3^5x_4x_5$ is the highest ordered term of $E_5\cdot E_3^2\cdot E_2^2$. If $h_1 = 1$, then $h_2 = h_3 = \cdots = h_k = 1$, where

$$h_1 + h_2 + \cdots + h_k = m.$$

This cannot be expressed as the highest ordered term of a product of elementary symmetric functions of weight less than m, because in every such product $h_1 \geqq 2$. Then $\Sigma x_1x_2\cdots x_m$ would be, according to the above definition, an elementary symmetric function. This is in accordance with the usual definition of elementary symmetric functions.

THEOREM. *Any polynomial of degree m symmetric in the n elements*

$$x_1, x_2, \cdots, x_n$$

is equal to an isobaric polynomial of weight m, with integral coefficients, in the $E_k^{(j)}$ and the coefficients of the polynomial.

The proof is based on induction on the order of the symmetric polynomials. The theorem need only be proved for simple symmetric polynomials, since other symmetric polynomials are sums of the simple symmetric polynomials. Given any simple symmetric polynomial, assume all simple symmetric polynomials of lower order capable of being expressed as polynomials in the $E_k^{(j)}$. Since the $E_k^{(j)}$ are defined so that they contain all simple symmetric polynomials which are not of the highest order in some product of $E_k^{(j)}$, every other simple symmetric polynomial is the highest ordered polynomial in some product $E_{i_1}^{(j_1)}\cdot E_{i_2}^{(j_2)}\cdots E_{i_k}^{(j_k)}$, or in other words,

$$S_m^{(j)}(x_n) = E_{i_1}^{(j_1)}\cdot E_{i_2}^{(j_2)}\cdots E_{i_k}^{(j_k)} - (\text{lower ordered terms}).$$

But lower ordered terms are expressible as polynomials in $E_k^{(j)}$; hence $S_m^{(j)}(x_n)$ is also expressible as a polynomial in the $E_k^{(j)}$. The theorem is true for the lowest ordered simple symmetric polynomials of every degree, because each such $\Sigma x_1x_2 \cdots x_k$ is defined as an $E_k^{(j)}$. The induction proof is then complete, and furthermore the coefficients are obviously integral, and the terms are isobaric of weight m.

In order to establish that the set of polynomials $E_k^{(j)}$, $k = 1, 2, \cdots, m$, form a fundamental set of order m, it remains to prove that every $S_m^{(j)}(x_n)$ can be represented uniquely as a polynomial in the $E_k^{(j)}$. Before this can be accomplished it must be proved that the highest ordered terms in products of $E_k^{(j)}$ are distinct. It has been proved that in the highest ordered term of a product the last letter of the left factor must be the same letter as the first in the right factor; it has also been proved that the exponents of the first and last letters of any $E_k^{(j)}$ are unity. Let $h_1 = x_{i_1} \cdots x_{i_k}$ be the highest ordered term in $E_i^{(j_i)}$ and let $h_{j \cdots r} = x_{i_k} \cdots x_{i_j}$ be the highest ordered term beginning with x_{i_k} in the product $E_j^{(j_j)} \cdots E_r^{(j_r)}$. Then

$$h_1 h_{j \cdots r} = [x_{i_1} \cdots x_{i_k}][x_{i_k} \cdots x_{i_j}]$$

is the highest ordered term in $E_i^{(j_i)} \cdot [E_j^{(j_j)} \cdots E_r^{(j_r)}]$. Now suppose it is also the highest ordered term in $E_l^{(j_l)} \cdot [E_m^{(j_m)} \cdots E_s^{(j_s)}]$, where $l > i$. Let

$$h_l = [x_{i_1} \cdots x_{i_k} \cdot x_{i_k} \cdots x_{i_n}]$$

be the highest ordered term in $E_l^{(j_l)}$ and let $h_{m \cdots s} = [x_{i_n} \cdots x_{i_j}]$ be the highest ordered term beginning with x_{i_n} in $E_m^{(j_m)} \cdots E_s^{(j_s)}$. Then

$$h_l h_{m \cdots s} = [x_{i_1} \cdots x_{i_k} \cdot x_{i_k} \cdots x_{i_n}][x_{i_n} \cdots x_{i_j}]$$

is the highest ordered term in $E_l^{(j_l)}[E_m^{(j_m)} \cdots E_s^{(j_s)}]$. But

$$h_{j \cdots r} = x_{i_k} \cdots x_{i_n} \cdot x_{i_n} \cdots x_{i_j}$$

is the highest ordered term beginning with x_{i_k} in $E_j^{(j_j)} \cdots E_r^{(j_r)}$, and $x_{i_k} \cdots x_{i_n}$ must be the highest ordered term beginning with x_{i_k} in some $S_m^{(j)}(x_n) = \Sigma x_{i_k} \cdots x_{i_n}$. By a previous theorem

$$S_m^{(j)}(x_n) = E_u^{(j_u)} E_v^{(j_v)} \cdots E_w^{(j_w)} - \text{(lower ordered terms)}.$$

Consequently the highest ordered term in $E_l^{(j_l)}$ is the highest ordered term in $E_i^{(j_i)} \cdot E_u^{(j_u)} E_v^{(j_v)} \cdots E_w^{(j_w)}$. This is a contradiction. If $[x_{i_1} \cdots x_{i_k}][x_{i_k} \cdots x_{i_j}]$ is the highest ordered term in $E_i^{(j_i)} \cdot E_j^{(j_j)} \cdots E_r^{(j_r)}$ and the highest ordered term also in $E_i^{(j_i)}[E_m^{(j_m)} \cdots E_s^{(j_s)}]$, where one can assume $E_j^{(j_j)} \neq E_m^{(j_m)}$, the first part of the proof can be reapplied to the second two factors $[E^{(j_j)} \cdots E^{(j_r)}]$ and $[E^{(j_m)} \cdots E^{(j_s)}]$, and a contradiction is reached.

THEOREM. *Every simple symmetric polynomial of degree m can be represented uniquely as a polynomial in the set of polynomials $E_k^{(j)}$, $k = 1, 2, \cdots, m$.*

Let $F(E_m^{(j)})$ be a polynomial in $E_1, \cdots, E_m^{(j)}$ and let one term be

$$P = ME_{i_1}^{(j_1)} E_{i_2}^{(j_2)} \cdots E_{i_k}^{(j_k)}$$

(all like terms being combined into one term P). If now the $E_k^{(j)}$ be replaced by the x_i's, a symmetric polynomial $\phi(x_i)$ of the x_i's is obtained with the highest term

$$A = Mx_{k_1}^{\nu_1}x_{k_2}^{\nu_2} \cdots x_{k_n}^{\nu_n}.$$

The terms P of $F(E_m^{(j)})$ are ordered so that P is of higher order than P' if after substitution of the x_i's the corresponding term A has a higher order than A'. For different terms P, the highest ordered terms A will be different by the ordering and the preceding theorem. Consequently the highest term P of $F(E_m^{(j)})$ has the same coefficient as the highest term A of $\phi(x_i)$. Hence if every coefficient of $F(E_m^{(j)})$ is not zero, every coefficient of $\phi(x_i)$ is not zero. If $\psi(x_i) = F(E_m^{(j)})$ and $\psi(x_i) = F_1(E_m^{(j)})$, where every coefficient of $F - F_1$ is not zero, upon substitution of the x_i's a contradiction is reached. Hence the representation of $S_m^{(j)}(x_n)$ in terms of the $E_k^{(j)}$ is unique. Consequently there can be no polynomial relationship among the $E_k^{(j)}$.

This completes the proof that the $E_k^{(j)}$, $k = 1, 2, \cdots, m$, form a fundamental set of order m.

5. **A second ordering and a second fundamental set.** The terms of a simple symmetric polynomial may be ordered as in the first order. The term $x_{i_1}^{\nu_1}x_{i_2}^{\nu_2} \cdots x_{i_k}^{\nu_k}$ is said to be of higher order than the term $x_{i'_1}^{\nu_1}x_{i'_2}^{\nu_2} \cdots x_{i'_k}^{\nu_k}$, if the first non-zero difference in the subscripts $i - i'$ is less than zero.

The simple symmetric polynomials of degree m are ordered in the following manner. Let $x_{j_1}^{\nu_1}x_{j_2}^{\nu_2} \cdots x_{j_k}^{\nu_k}$ be the highest ordered term of $S_m^{(i)}(x_n)$ and $x_{j'_1}^{\nu'_1}x_{j'_2}^{\nu'_2} \cdots x_{j'_k}^{\nu'_k}$ be the highest ordered term of $S_m^{(j)}(x_n)$. Write these highest ordered terms in the form $x_{i_1}x_{i_2} \cdots x_{i_m}$ and $x_{i'_1}x_{i'_2} \cdots x_{i'_m}$ so that every exponent is unity. The simple symmetric polynomials may be ordered according to the order of the highest ordered terms. $x_{i_1}x_{i_2} \cdots x_{i_m}$ is said to be of higher order than $x_{i'_1}x_{i'_2} \cdots x_{i'_m}$ if $i_1 < i'_1$. If $i_1 = i'_1$, then $x_{i_1}x_{i_2} \cdots x_{i_m}$ is of higher order if $i_2 < i'_2$, and if $i_1 = i'_1$, $i_2 = i'_2$, then $x_{i_1}x_{i_2} \cdots x_{i_m}$ is of higher order if $i_3 < i'_3$, etc. This process is continued until all simple symmetric polynomials of degree m are ordered. The polynomials of degree m are considered to be of higher order than all those of lower degree.

Under this ordering it can be proved that the highest ordered term of the product of two simple symmetric polynomials $S_m^{(i)}(x_k)S_n^{(j)}(x_k)$ is equal to the highest ordered term of $S_m^{(i)}(x_k)$ multiplied on the right by the highest ordered term of $S_n^{(j)}(x_k)$.

Define a set of functions as follows. (a) $H_1 = \Sigma x_1$. (b) Every simple symmetric polynomial of degree m of which the highest ordered term is not the highest ordered term of some product of weight m of $H_n^{(j)}$, $n = 1, 2, \cdots, m-1$ and $j = 1, 2, \cdots, j_n$, is to be defined as an $H_m^{(j)}$, $j = 1, 2, \cdots, j_m$.

All $H_m^{(j)}$ through the fourth degree are in the following set:

$$H_1 = \Sigma x_1,$$
$$H_2 = \Sigma x_1 x_2,$$
$$H_3^{(1)} = \Sigma x_1 x_2^2, \qquad H_3^{(2)} = \Sigma x_1 x_2 x_3,$$
$$H_4^{(1)} = \Sigma x_1 x_2 x_1 x_3, \qquad H_4^{(2)} = \Sigma x_1 x_2^3, \qquad H_4^{(3)} = \Sigma x_1 x_2^2 x_3,$$
$$H_4^{(4)} = \Sigma x_1 x_2 x_3 x_2, \qquad H_4^{(5)} = \Sigma x_1 x_2 x_3^2, \qquad H_4^{(6)} = \Sigma x_1 x_2 x_3 x_4.$$

It can be proved that the above set forms a fundamental set which is also a direct generalization of the elementary symmetric functions. The proofs are almost identical with those of the preceding paragraphs.

6. **General properties of fundamental sets.** The choice of simple symmetric functions which can be used for fundamental polynomials of order m is not unique, but there are several properties which are common to all, which can now be stated and proved. Clearly, every fundamental set of order m must have Σx_1 as the first polynomial. It can be proved that every simple symmetric function of degree m is contained in $(\Sigma x_1)^m$. The statement is easily verified for degrees one and two, and the proof can be completed by induction. Assume $(\Sigma x_1)^{m-1}$ contains every simple symmetric polynomial of degree $m - 1$. Multiply $(\Sigma x_1)^{m-1}(\Sigma x_1) = (\Sigma x_1)^m$. The m-th position is filled in every possible way, since Σx_1 contains every x_i. Hence the above conclusion is true.

THEOREM. *The simple symmetric polynomials of degree m which involve two elements x_1, x_2 require for representation as polynomials in a fundamental set of order m one and only one fundamental polynomial of every degree* $1, 2, \cdots, m$.

The theorem is proved by induction. It can be demonstrated for degrees one and two that

$$F_1 = x_1 + x_2, \qquad F_2 = x_1 x_2 + x_2 x_1, \qquad (F_1)^2 - F_2 = x_1^2 + x_2^2,$$

or

$$F_1' = x_1 + x_2, \qquad F_2' = x_1^2 + x_2^2, \qquad (F_1')^2 - F_2 = x_1 x_2 + x_2 x_1.$$

Assume that all simple symmetric polynomials through degree $m - 1$ can be expressed as polynomials in a set of fundamental polynomials $F_1, F_2, \cdots, F_{m-1}$, where there is one and only one F_k of each weight $k \leqq m - 1$. There is a totality of 2^{m-1} simple symmetric polynomials of degree m involving two letters, since m positions can be filled by two letters in 2^m ways, but those two permutations belong to one simple symmetric polynomial which differ only in the interchange of x_1 and x_2. From the above assumption the total number of products of the F_k of weight m is $2^{m-1} - 1$ because the total number of compositions[1] of m from $1, 2, \cdots, m$ is 2^{m-1}, and if m is excluded, the number is $2^{m-1} - 1$. The representation of the simple symmetric polynomials of degree m, the $S_m^{(i)}(x_n)$,

[1] Macmahon, *Combinatorial Analysis*, vol. I, p. 151.

as polynomials in the F_k must be unique; hence the F_k must be so chosen that a polynomial $P_1(F_1, F_2, \cdots, F_m)$ can equal another polynomial

$$P_2(F_1, F_2, \cdots, F_m),$$

if and only if the coefficients of like terms are equal. Consider the products of the F_k of weight m as $2^{m-1} - 1$ equations in the 2^{m-1} unknowns $S_m^{(j)}(x_n)$. It has been proved that every $S_m^{(j)}(x_n)$ is present in at least one product, namely, $(\Sigma x_1)^m$, and it has been proved that a product of simple symmetric polynomials, and therefore a product of the F_k, is a sum of simple symmetric polynomials with coefficients positive unity. The rank of the matrix of the coefficients of the $S_m^{(j)}(x_n)$ is equal to $2^{m-1} - 1$, which is the number of equations. That is, the rank is equal to the number of products of the F_k which are of weight m, because there is no polynomial relationship between the products. Since there are 2^{m-1} unknown $S_m^{(j)}(x_n)$, it is necessary that one $S_m^{(j)}(x_n)$ be assigned arbitrarily so that one can solve for the other $S_m^{(j)}(x_n)$ in terms of the products of the F_k, $k = 1, 2, \cdots, m-1$, of weight m and the fixed $S_m^{(j)}(x_n)$. Let this $S_m^{(j)}(x_n) = F_m$.

From this theorem it is clear that the fundamental polynomials of order m form an infinite sequence as m increases indefinitely.

THEOREM. *For every fundamental set of order m the number of polynomials for each degree $n \leqq m$ is unique, and is equal to the total number of simple symmetric polynomials of degree n minus the number of products of weight n of the fundamental set $F_k^{(j)}$, $k = 1, 2, \cdots, n-1$.*

Let the number of products of weight n of the $F_k^{(j)}$ be t. Let the number of $S_n^{(j)}(x_n)$ of degree n be s. The t products are t equations in the s unknowns $S_n^{(j)}(x_n)$ as in the preceding theorem. Since there is no polynomial relationship between the t products, the rank of the matrix of the coefficients of the $S_n^{(j)}(x_n)$ is t. There is then a solution in which the $S_n^{(j)}(x_n)$ are expressed as polynomials of the $F_k^{(j)}$, $k = 1, 2, \cdots, n-1$, and $s-t$ of the $S_n^{(j)}(x_n)$. Assign these $s-t$ of the $S_n^{(j)}(x_n)$ as $F_m^{(j)}$, $j = 1, 2, \cdots, s-t$. The preceding theorem necessitates that $t < s$, because at least one $F_m^{(j)}$ exists for every m.

7. **Remarks.** As the degree m increases, the number of fundamental polynomials increases by a finite number for each m. This distinguishes them from elementary symmetric functions. Because of this difference it cannot be expected that all the theorems on symmetric functions of commutative elements can be generalized to non-commutative elements. For example, the theorem which states that every symmetric polynomial can be expressed as a polynomial of the $S_k = \Sigma x_1^k$, which are sums of like powers of the elements, does not hold in the case of non-commutative elements. Another theorem which does not hold is the following.

A symmetric polynomial in $x_1, x_2, \cdots, x_n$ when written in terms of the elementary symmetric functions $E_1, E_2, \cdots, E_m$ will be of the same degree in the E's as it was in any one of the x's.

An example to show this is not true is given by $E_4^{(6)} = \Sigma x_1 x_2^2 x_3$, of first degree in $E_4^{(6)}$, but of second in x_2.

It has been proved, in general, that every simple symmetric polynomial can be expressed with integral coefficients as a polynomial in the fundamental polynomials called $E_m^{(j)}$. It can be exhibited that all simple symmetric polynomials in the five non-commutative elements x_1, x_2, x_3, x_4, x_5 of degrees 1, 2, 3, 4, 5 can be expressed as polynomials in the $H_m^{(j)}$ with coefficients which are not only integral but are positive or negative unity. It is conjectured that every simple symmetric polynomial of degree m can be expressed as a polynomial in the fundamental polynomials with coefficients equal to positive or negative unity. For all the specific cases computed it was found that every simple symmetric polynomial of degree m when expressed as a polynomial in a fundamental set includes as a term an $F_m^{(j)}$ of weight m for at least one value of j. These and several other properties of fundamental sets are open for consideration.

UNIVERSITY OF WISCONSIN.

FUNCTIONS ARISING FROM DIFFERENTIAL EQUATIONS AND SERVING TO GENERALIZE A THEOREM OF LANDAU AND CARATHÉODORY

By John W. Cell

Introduction. The hypergeometric linear differential equation which has for solutions the quarter periods of elliptic functions has been studied extensively by Fuchs[1] and Tannery.[2] By the use of a particular quotient of two of its solutions Picard and Landau[3] proved their remarkable theorems on analytic functions. If a certain transformation is made so that the exponents at the singular points $(0, 1, \infty)$ of this hypergeometric equation are all equal to each other, the equation so obtained is invariant with respect to the linear fractional dihedral group of order six, generated by $z' = 1 - z$ and $z' = 1/z$, where z is a complex variable.

The cyclic, dihedral, tetrahedral, octahedral and icosahedral groups are the only groups of finite order which are representable on linear fractional substitutions of a complex variable.[4] We shall specialize the exponent differences at the three singular points of the hypergeometric equation and then make the substitution $z = x^n$ on the independent variable. For each of the four specializations to be made we shall obtain an equation which is invariant with respect to some one of the first four groups named above, which is such that the exponent difference at each singular point is zero, and which has the property that an appropriate quotient function of two of its solutions has properties quite similar to those of the quotient function already mentioned.

We shall thus obtain four quotient functions, and by their use we shall obtain specific formulas for the radius of the circle in which every function of the form $F(x) = a_0 + a_1x + a_2x^2 + \cdots$ $(a_1 \neq 0)$ must either have a singularity or assume one of a certain set of values as, for example, the n n-th roots of unity. Moreover, these radii will depend only on a_0, a_1, and this set of values.

1. **Specializations of the hypergeometric equation.** In the hypergeometric equation

$$(1)\quad 4z(z-1)v'' + 4\{(z-1)(1-\lambda) + z(1-\mu)\}v' + \{(1-\lambda-\mu)^2 - \nu^2\}\, v = 0$$

the singular points are at $z = 0$, 1, and ∞ with exponents $0, \lambda$; $0, \mu$; $\frac{1}{2}(1-\lambda-\mu-\nu)$, $\frac{1}{2}(1-\lambda-\mu+\nu)$, respectively.

Received August 26, 1935.

[1] L. Fuchs, Journal für Mathematik, vol. 71 (1870), pp. 91–127.

[2] J. Tannery, Annales de l'Ecole Normale, (2), vol. 15 (1879), pp. 169–194.

[3] E. Landau, Vierteljahrschrift der Natur. Gesellschaft, vol. 51 (1906), pp. 252–318.

[4] F. Klein, *Lectures on the Icosahedron*, p. 126.

If we set $\lambda = \mu = \nu = 0$ and denote by $V(z)$ the solution which is regular in the neighborhood of the origin and which is such that $V(0) = 1$, then $V(1 - z)$ is also a solution of equation (1) and is regular in the neighborhood of $z = 1$. We define

$$T(z) = iV(1 - z)/V(z).$$

This is the quotient function which was used by Picard and Landau. $T(z)$ *possesses the following properties:*[5] *if* z *stays away from* 0, 1 *and* ∞, *then* (1) $T(z)$ *is regular,* (2) $\Im\,[T(z)] > 0$, (3) $T'(z) \neq 0$; *if* $z = z(T)$ *is the inverse function and if* $\Im\,(T) > 0$, *then* (4) $z(T)$ *is single-valued and regular,* (5) $z(T)$ *stays away from* 0 *and* 1; *moreover,* (6) *the axis of reals is a natural boundary for* $z(T)$, *and* (7) $z(T)$ *is a fuchsian function.*

In a linear differential equation of the second order whose singular points are all regular singular points, the inverse of the quotient of two solutions is a single-valued function if and only if the exponent difference at each singular point is individually either zero, the reciprocal of an integer greater than 1, or 1 with the condition that a transformation can be made so that this singular point becomes an ordinary point in the resulting differential equation.[6] Moreover, the inverse function of the quotient of two such linearly independent solutions will stay away from those values which correspond to singular points with exponent difference zero.[7]

We make the transformation $z = f(x)$, $v(z) = u(x)$ on equation (1) and require $f(x)$ to be single-valued and such that the resulting second order differential equation has the following properties: (A) the coefficients are rational functions, (B) the singular points are all regular singular points, (C) the exponent difference at each singular point is either zero or the reciprocal of a positive integer. (If this integer is 1, a further condition is necessary as before.) Moreover, λ, μ and ν are to be either zero or the reciprocal of an integer greater than 1. A straightforward computation shows that if n is a positive integer and c is a non-zero constant, only the following six transformations possess the requisite properties: $z = cx^n$, $z = 1 - cx^n$, $z = c/x^n$, $z = 1/(1 - cx^n)$, $z = (cx^n - 1)/(cx^n)$, $z = cx^n/(cx^n - 1)$. But these are all essentially $z = cx^n$, since we may obtain the others by first permuting the singular points of the hypergeometric equation and then making this transformation. We shall work out the case for $c = 1$, since the general case is obtainable from this by a transformation of the form $x' = kx$, where $k \neq 0$.

We make the transformation $z = x^n$, $v(z) = u(x)$ on equation (1) and obtain

$$\begin{aligned} &4x(x^n - 1)u'' + 4\{(1 - \lambda n)(x^n - 1) + n(1 - \mu)x^n\}u' \\ &\quad + n^2\{(1 - \lambda - \mu)^2 - \nu^2\}x^{n-1}u = 0. \end{aligned} \tag{2}$$

[5] Tannery, loc. cit.; Klein, *Vorlesungen über die Hypergeometrische Funktion*, 1933, p. 291.
[6] L. R. Ford, *Automorphic Functions*, pp. 298 and 304.
[7] Klein, see footnote 5, p. 292; H. Poincaré, *Oeuvres*, vol. 2, p. 14.

If $\sigma = e^{\pi i/n}$, the exponent differences at the singular points of equation (2) are respectively: λn at $x = 0$, μ at $x = 1, \sigma, \sigma^2, \cdots, \sigma^{n-1}$, and νn at $x = \infty$. If the quotient of two solutions of equation (2) is to have properties similar to those of $T(z)$, we need to consider only the following cases:

Case I. $\lambda = 1/n$, $\mu = \nu = 0$,

Case II. $\lambda = \nu = 1/n$, $\mu = 0$,

Case III. $\lambda = \mu = \nu = 0$.

The case of $\lambda = \mu = 0$, $\nu = 1/n$ is a transformation of case I by the transformation $x' = 1/x$. We shall study the three cases in §§2–4.

For convenience we introduce the following notation. We use $F(a, b, c, x)$ to denote the hypergeometric series

$$1 + \frac{a \cdot b}{1 \cdot c} x + \frac{a(a+1)(b)(b+1)}{1 \cdot 2 \cdot c(c+1)} x^2 + \cdots$$

and its analytic continuations. In each of the three cases to be considered we cut the plane by joining each singular point of the corresponding differential equation to $x = \infty$ by straight lines which, if continued, pass through the origin. If $x = 0$ is a singular point, we join it to $x = \infty$ along the negative axis of reals. We use the term "principal determination" to describe the function defined by the hypergeometric series and its analytic continuations in this cut plane and denote the function, so defined, by $F^*(a, b, c, x)$. Similar definitions will pertain to the other solutions of equation (2) which are to be considered. We denote by $D(a, b, c, x)$ the function defined for $|x| < 1$ by

$$\lim_{\epsilon = 0} \epsilon^{-1}\{F(a + \epsilon, b + \epsilon, c + 2\epsilon, x) - e^{-\epsilon \log x} F(a, b, c, x)\},$$

(where the principal determination is assigned to $\log x$) and its analytic continuations. We observe that $D^*(a, b, c, x)$, which is this function in the cut plane, is expansible in a series whose general term is the corresponding $(j + 1)$-th term in the hypergeometric series multiplied by

$$\{\psi(a + j) + \psi(b + j) - 2\psi(c + j) - \psi(a) - \psi(b) + 2\psi(c) + \log x\},$$

where $\psi(a)$ is the classic function $\Gamma'(a)/\Gamma(a)$.

2. **Case** I. $\lambda = 1/n$, $\mu = \nu = 0$. The hypergeometric equation (1) for this case has the fundamental system of solutions $F(\alpha, \alpha, 2\alpha, z)$ and $e^{(\log z)/n} F(1 - \alpha, 1 - \alpha, 2 - 2\alpha, z)$, where $\alpha = (n - 1)/(2n)$. Moreover, equation (2) for this case becomes

$$4(x^n - 1)u'' + 4nx^{n-1} u' + (n - 1)^2 x^{n-2} u = 0. \tag{3}$$

The results for $n = 1$ are trivial, so we shall suppose that $n \geqq 2$. We substitute $u(x) = (1 - x^n)^{(1-n)/(2n+2)} y(x)$ and obtain a differential equation which has regular singular points at the n n-th roots of unity and at ∞ with equal exponents $(n - 1)/(2n + 2)$, and which is invariant with respect to the cyclic group of order n, generated by $x' = \sigma x$. For the special case $n = 3$ it is invariant

with respect to the tetrahedral group of order 12, generated by $x' = \sigma x$ and $x' = (x + 2\sigma^2)/(\sigma x - 1)$. Equation (3) for this case is also invariant with respect to this same cyclic group.

We refer to the identities connecting the several solutions of the hypergeometric equation,[8] specialize them by using standard limiting processes, make our substitution $z = x^n$, and obtain the following four identities which are valid for x in the sector S_1 bounded by the rays $(0, 1, \infty)$ and $(0, \sigma, \infty)$:

$$F^*(\alpha, \alpha, 2\alpha, x^n) = \frac{\Gamma(2\alpha)}{\Gamma^2(\alpha)} \{[2\psi(1) - 2\psi(\alpha)] F^*(\alpha, \alpha, 1, 1 - x^n) - D^*(\alpha, \alpha, 1, 1 - x^n)\},$$

$$xF^*(1 - \alpha, 1 - \alpha, 2 - 2\alpha, x^n) = \frac{\Gamma(2 - 2\alpha)}{\Gamma^2(1 - \alpha)} \{[2\psi(1) - 2\psi(1 - \alpha)] F^*(\alpha, \alpha, 1, 1 - x^n) - D^*(\alpha, \alpha, 1, 1 - x^n)\}, \tag{4}$$

$$F^*(\alpha, \alpha, 2\alpha, x^n) = \frac{\Gamma(2\alpha)}{\Gamma^2(\alpha)} e^{\alpha(\pi i - n \log x)} \{[2\psi(1) - 2\psi(\alpha) - \pi i] F^*(\alpha, 1 - \alpha, 1, x^{-n}) - D^*(\alpha, 1 - \alpha, 1, x^{-n})\},$$

$$xF^*(1 - \alpha, 1 - \alpha, 2 - 2\alpha, x^n) = \frac{\Gamma(2 - 2\alpha)}{\Gamma^2(1 - \alpha)} e^{\pi i(1-\alpha) - \alpha n \log x}\{[2\psi(1) - 2\psi(1 - \alpha) - \pi i] F^*(\alpha, 1 - \alpha, 1, x^{-n}) - D^*(\alpha, 1 - \alpha, 1, x^{-n})\}.$$

We observe that the hypergeometric series in the definition of $F^*(\alpha, \alpha, 2\alpha, x^n)$ converges for $|x| < 1$, $F^*(\alpha, 1 - \alpha, 1, x^{-n})$ for $|x| > 1$, and $F^*(\alpha, \alpha, 1, 1 - x^n)$ for $|1 - x^n| < 1$ and hence in the interior of the rose $\rho = 2 \cos n\theta$ (in polar coördinates with pole at the origin). We use $F^*(\alpha, \alpha, 1, 1 - x^n)$ to denote the function defined by the corresponding hypergeometric series for x in the leaf of the rose about $x = 1$ and its analytic continuations in the cut plane. Corresponding remarks apply to the other three solutions which occur in these identities. Other identities, valid for x in the other $n - 1$ sectors of the plane, are easily obtainable from the group invariance property. These identities make evident the fact that the hypergeometric series in the definition of $F^*(\alpha, \alpha, 1, 1 - x^n)$ defines n distinct functions.

THEOREM 1. *The quotient function*

$$t_1(x) = K_1 x F(1 - \alpha, 1 - \alpha, 2 - 2\alpha, x^n)/F(\alpha, \alpha, 2\alpha, x^n),$$

where $K_1 = \Gamma(2\alpha)\Gamma^2(1 - \alpha)/\Gamma(2 - 2\alpha)\Gamma^2(\alpha)$, *has the following properties: if* $x^n \neq 1$ *and* $x \neq \infty$, *then* (1) $t_1(x)$ *is regular,* (2) $|t_1(x)| < 1$, (3) $t_1'(x) \neq 0$; *if* $x = x(t)$ *is the inverse function, for* $|t| < 1$, (4) $x(t)$ *is single-valued and regular,* (5) $\{x(t)\}^n \neq 1$; *moreover,* (6) $|t| = 1$ *is a natural boundary for* $x(t)$, *and* (7) $x(t)$ *is a fuchsian function.*

[8] E. Goursat, Annales de l'Ecole Normale, (2), vol. 17 (1881), appendix; especially pages 20–21, 28–30, 34–37.

We shall first establish the second proposition as follows. Because of the manner in which the two solutions in the definition of $t_1(x)$ have been defined, we may consider $N(x^n) = t_1(x) = N(z)$. But $N^*(z)$ maps the upper half z-plane upon a circular arc triangle with vertices $N = 0$ (corresponding to $z = 0$), $N = 1$ ($z = 1$) and $N = e^{\pi i/n}$ ($z = \infty$) and with interior angles π/n, 0, 0, respectively.[9] Then $t_1^*(x)$ maps the sector S_1 upon this same circular arc triangle. We use the principle of reflection[10] to see that $t_1^*(x)$ maps the cut x-plane upon the circular arc $2n$-gon R which is such that if we join each vertex to the origin the $2n$ triangles so formed are all congruent (with respect to the powers of the substitution $t' = \sigma t$) to the circular arc triangle described above or to its reflection in the axis of reals.

We define $f(x)]_{x=a}$ as the result of taking $f(x)$ about a simple closed contour which encircles the singular point $x = a$ in a counterclockwise direction and which encircles no other singular point of $f(x)$.

We make use of the identities (4) and similar identities to show that $t_1(x)$ possesses the following circuit properties:

$$t_1(x)]_{x=\sigma^k} = \frac{\{1 + i \cot(\pi/2n)\}t_1(x) - i\sigma^k \cot(\pi/2n)}{i\sigma^{-k} \cot(\pi/2n)t_1(x) + \{1 - i \cot(\pi/2n)\}}, \tag{5}$$

$$(k = 0, 1, 2, \cdots, n-1).$$

To obtain the general map, we apply to the region R the group of substitutions G generated by the substitutions (5) and their inverses. We make use of the idea of isometric circles[11] of these substitutions to see that the map so obtained is interior to $|t| = 1$ except for the vertices of R and points congruent to these vertices (these points are the images of the singular points $1, \sigma, \sigma^2, \cdots, \sigma^{n-1}, \infty$) which all lie on the above circle. Each vertex and the points congruent to it form an everywhere dense set of points on this circle.

Proposition (1) of our theorem can now be established from the observation that the numerator and denominator of $t_1(x)$ together form a fundamental system of solutions of equation (3), and hence $t_1(x)$ is regular for x away from the singular points of that equation, except possibly for simple poles. But from the map already described we see that $t_1(x)$ has no poles. As a corollary we observe that since these two solutions cannot vanish simultaneously for x away from the singular points of (3), the denominator of $t_1(x)$ and hence $F(\alpha, \alpha, 2\alpha, x^n)$ has no zeros and no poles for x away from these singular points. But

$$t_1'(x) = K_1(1 - x^n)^{-1} [F(\alpha, \alpha, 2\alpha, x^n)]^{-2},$$

and hence proposition (3) of the theorem is evident.

Proposition (4) follows from the fact that the general map fills up, without overlapping, the interior of $|t| = 1$. From our observations concerning the

[9] See Klein, *Hypergeometrische Funktion*, p. 196; Ford, loc. cit., pp. 304–305.

[10] See Bieberbach, *Lehrbuch der Funktionentheorie*, (3rd ed.) vol. 1, p. 225.

[11] Ford, loc. cit., pp. 26–29 and Chapter III.

images of the singular points of equation (3), the propositions (5) and (6) follow. $x(t)$ is a fuchsian function because it is automorphic with respect to the group G, which has $|t| = 1$ as its fixed circle. Hence the theorem is established. We observe that these properties may also be established from a direct study of equation (3) without reference to its relation to the hypergeometric equation.

Let $w = e^{2\pi i/3}$. We define

$$g_k(x) = \frac{-wt_k(x) + w}{t_k(x) - w^2} \tag{6}$$

for $k = 1$ here and for other k later as we introduce new functions $t_k(x)$. Then $g_1(x)$ has properties similar to those of $t_1(x)$ except that $|t| \leqq 1$ becomes $\Im(g) \geqq 0$.

We make the transformation $x' = 1/x$ and obtain the following

COROLLARY. *The quotient function $t_2(x) = t_1(1/x)$ has the following properties: if $x^n \neq 1$ and $x \neq 0$, then (1) $t_2(x)$ is regular, (2) $|t_2(x)| < 1$, (3) $t_2'(x) \neq 0$; if $x = x(t)$ is the inverse function and if $|t| < 1$, then (4) $x(t)$ is single-valued and regular except for simple poles, (5) $x(t) \neq 0$ and $\{x(t)\}^n \neq 1$; moreover, (6) $|t| = 1$ is a natural boundary for $x(t)$, and (7) $x(t)$ is a fuchsian function.*

3. **Case** II. $\lambda = \nu = 1/n$, $\mu = 0$. Equation (2) for this case becomes

$$4(x^n - 1)u'' + 4nx^{n-1}u' + n(n-2)x^{n-2}u = 0. \tag{7}$$

We obtain trivial results if $n = 1$ or $n = 2$, so we shall suppose that $n \geqq 3$. If we make the transformation $u(x) = (1 - x^n)^{(2-n)/(2n)} y(x)$, we obtain an equation which is invariant with respect to the dihedral group of order $2n$, generated by $x' = \sigma x$ and $x' = 1/x$.

As in the preceding section we obtain the following identities, valid for x in S_1, with $\alpha = 1/2$, $\beta = (n-2)/(2n)$ and conventions as before for the solutions involved:

$$F^*(\alpha, \beta, \alpha + \beta, x^n) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\{[2\psi(1) - \psi(\alpha) - \psi(\beta)]F^*(\alpha, \beta, 1, 1 - x^n)$$
$$- D^*(\alpha, \beta, 1, 1 - x^n)\},$$

$$xF^*(1-\alpha, 1-\beta, 2-\alpha-\beta, x^n) = \frac{\Gamma(2-\alpha-\beta)}{\Gamma(1-\alpha)\Gamma(1-\beta)}\{[2\psi(1) - \psi(1-\alpha)$$
$$- \psi(1-\beta)]F^*(\alpha, \beta, 1, 1 - x^n) - D^*(\alpha, \beta, 1, 1 - x^n)\}, \tag{8}$$

$$F^*(\alpha, \beta, \alpha + \beta, x^n) = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+\beta-1)}{\Gamma^2(\beta)} e^{\alpha(\pi i - n\log x)} F^*(1-\alpha, 1-\beta,$$
$$2 - \alpha - \beta, x^{-n}) + \frac{\Gamma(\alpha+\beta)\Gamma(1-\alpha-\beta)}{\Gamma^2(\alpha)} e^{\beta(\pi i - n\log x)} F^*(\alpha, \beta, \alpha+\beta, x^{-n}),$$

$$xF^*(1-\alpha, 1-\beta, 2-\alpha-\beta, x^n) = e^{\pi i/n}\left\{\frac{\Gamma(2-\alpha-\beta)\Gamma(\alpha+\beta-1)}{\Gamma^2(1-\alpha)} e^{\alpha(\pi i - n\log x)}\right.$$

$$F^*(1-\alpha, 1-\beta, 2-\alpha-\beta, x^{-n}) + \frac{\Gamma(2-\alpha-\beta)\Gamma(1-\alpha-\beta)}{\Gamma^2(1-\beta)}$$

$$\left. e^{\beta(\pi i - n\log x)} F^*(\alpha, \beta, \alpha+\beta, x^{-n})\right\}.$$

THEOREM 2. *The quotient function*

$$t_3(x) = K_3 xF(1-\alpha, 1-\beta, 2-\alpha-\beta, x^n)/F(\alpha, \beta, \alpha+\beta, x^n),$$

where $K_3 = \Gamma(\alpha+\beta)\Gamma(1-\beta)/\Gamma(2-\alpha-\beta)\Gamma(\beta)$ *has the following properties: if* $x^n \neq 1$, *then* (1) $t_3(x)$ *is regular,* (2) $|t_3(x)| < 1$, (3) $t_3'(x) \neq 0$; *if* $x = x(t)$ *is the inverse function and if* $|t| < 1$, *then* (4) $x(t)$ *is single-valued and regular except for simple poles,* (5) $\{x(t)\}^n \neq 1$; *moreover,* (6) $|t| = 1$ *is a natural boundary for* $x(t)$, *and* (7) $x(t)$ *is a fuchsian function.*

The proof of this theorem is so similar to the proof of the preceding theorem that we shall give only the following relations which are necessary in this proof:

$$(9)\qquad \begin{gathered} t_3(x)]_{x=\sigma^K} = \frac{(1+2i\cot(\pi/n))t_3(x) - 2i\sigma^K\cot(\pi/n)}{2i\sigma^{-K}\cot(\pi/n)t_3(x) + (1-2i\cot(\pi/n))}, \\ K = 0, 1, \cdots, n-1, \\ t_3'(x) = K_3(1-x^n)^{-1}[F(\alpha, \beta, \alpha+\beta, x^n)]^{-2}. \end{gathered}$$

$t_3(x)$ maps the cut plane upon the circular arc $2n$-gon which may be described thus: its vertices are at $t = 1$, $t = e^{\pi i/n}\cos(\pi/n)$, $t = e^{-\pi i/n}\cos(\pi/n)$ and the points congruent to these three points with respect to the powers of $t' = \sigma t$. If we join the first two points by a circular arc which makes a zero angle at $t = 1$ with the line segment joining $t = 1$ to the origin and which makes an angle π/n at the other point with the line segment joining this point to the origin, the other arcs of this $2n$-gon are congruent either to this arc or to its image in the axis of reals. The general map, as before, is obtained by the use of the isometric circle of the substitutions (9) and their inverses.

We have here the additional relation, valid for x in S_1,

$$t_3(x) = \frac{\{1 - i\cot(\pi/n)\}t_3(1/x) + i\cot(\pi/n)}{-i\cot(\pi/n)t_3(1/x) + \{1 + i\cot(\pi/n)\}},$$

where $t_3(1/x) = K_3F(1-\alpha, 1-\beta, 2-\alpha-\beta, x^{-n})/xF(\alpha, \beta, \alpha+\beta, x^{-n})$.

4. Case III. $\lambda = \mu = \nu = 0$. Equation (2) becomes

$$(10)\qquad 4x(x^n-1)u'' + 4\{(n+1)x^n - 1\}u' + n^2x^{n-1}u = 0.$$

If we make the transformation $u(x) = (x - x^{n+1})^{-n/(2n+4)}y(x)$, we obtain an equation which is invariant with respect to the dihedral group of order $2n$, generated by $x' = \sigma x$, $x' = 1/x$. For the particular case $n = 4$ it is invariant

with respect to the octahedral group of order 24, generated by $x' = \sigma x$, $x' = (x + 1)/(x - 1)$.

In this case, where we use the principal determination for the logarithm and where $F^*(\frac{1}{2}, \frac{1}{2}, 1, x^n) = f^*(x^n)$ and $D^*(\frac{1}{2}, \frac{1}{2}, 1, x^n) = d^*(x^n)$, the four identities, valid for x in S_1, become

$$
\begin{aligned}
&\pi f^*(x^n) = 4 \log 2 f^*(1 - x^n) - d^*(1 - x^n),\\
&\pi d^*(x^n) = (16 \log^2 2 - \pi^2) f^*(1 - x^n) - 4 \log 2 d^*(1 - x^n),\\
&\pi f^*(x^n) e^{\frac{1}{2} \log x} = (4i \log 2 + \pi) f^*(x^{-n}) - i d^*(x^{-n}),\\
&\pi d^*(x^n) e^{\frac{1}{2} \log x} = 16 i \log^2 2\, f^*(x^{-n}) + (\pi - 4i \log 2) d^*(x^{-n}).
\end{aligned} \tag{11}
$$

Theorem 3. *The quotient function*

$$g_4(x) = iF(\tfrac{1}{2}, \tfrac{1}{2}, 1, 1 - x^n)/F(\tfrac{1}{2}, \tfrac{1}{2}, 1, x^n)$$

has the following properties: if $x \neq 0$, $x^n \neq 1$, $x \neq \infty$, then (1) *$g_4(x)$ is regular,* (2) $\mathfrak{I}(g_4(x)) > 0$, (3) $g_4'(x) \neq 0$; *if $x = x(g)$ is the inverse function and if $\mathfrak{I}(g) > 0$, then* (4) *$x(g)$ is single-valued and regular,* (5) *$x(g) \neq 0$ and $\{x(g)\}^n \neq 1$; moreover,* (6) *the axis of reals is a natural boundary for $x(g)$, and* (7) *$x(g)$ is a fuchsian function.*

The proof of this theorem may be accomplished by the same means as before. For that purpose we need the relations

$$
\begin{aligned}
&g_4(x)]_{x=0} = g_4(x) + 2n,\\
&g_4(x)]_{x=\sigma^k} = \frac{(4k + 1)g_4(x) + 8k^2}{-2g_4(x) - (4k - 1)} \qquad (k = 0, 1, 2, \cdots, n - 1),\\
&g_4'(x) = n[\pi i x(1 - x^n)]^{-1}[F(\tfrac{1}{2}, \tfrac{1}{2}, 1, x^n)]^{-2}.
\end{aligned} \tag{12}
$$

The theorem may easily be established, however, by using the properties of the function $T(z)$ defined at the beginning of the first section. The propositions are obtainable by the direct substitution of $z = x^n$ in $T(z)$. The general map is essentially the map of the cut z-plane by $T(z)$[12] except that it here takes $2n$ of the triangular regions to constitute a fundamental region for the group generated by the substitutions (12a) and (12b) and their inverses.

5. **Extensions of the Landau theorem.** Let a_0 and $a_1 \neq 0$ be two given constants. In this section we shall use $F(x)$ to denote any function which is regular in the neighborhood of the origin and which has there the form $F(x) = a_0 + a_1 x + a_2 x^2 + \cdots$.

Landau's theorem[13] states that there exists a number depending only on a_0 and a_1, say $R(a_0, a_1)$, such that $F(x)$ has in or on the circle $|x| = R$ either a singularity, a zero, or assumes the value 1. Carathéodory[13] showed that the

[12] See Klein-Fricke, *Theorie der elliptischen Modulfunktionen*, vol. 1, p. 273.

[13] Landau, footnote 3.

least possible value of $R(a_0, a_1)$ for this theorem is $\phi(a_0, a_1) = 0$ if $a_0 = 0$ or $a_0 = 1$; otherwise

$$\phi(a_0, a_1) = \frac{2\Im(T(a_0))}{|a_1| \cdot |T'(a_0)|},$$

where $T(z)$ is an arbitrary branch of the quotient function defined in §1 of this paper. Landau's theorem implies the existence of a number M depending only on $a_0, a_1, \alpha_1, \alpha_2, \cdots, \alpha_k$, where α_i are distinct finite numbers and $k \geqq 2$, such that $F(x)$ in or on the circle $|x| = M$ either has a singularity or assumes a value α_i.

Let $g = g(x)$ be a general symbol for any function having the following properties: if $x = \alpha_i$ $(i = 1, 2, \cdots, m)$, where $m \geqq 3$ and the α_i are distinct finite numbers except in the case of α_m which may be ∞, then (1) $g(x)$ is regular, (2) $\Im(g(x)) > 0$, (3) $g'(x) \neq 0$; if $x = h(g)$ is the inverse function and if $\Im(g) > 0$, then (4) $h(g)$ is single-valued and regular except for simple poles [if $\alpha_m = \infty$, then $h(g)$ is to have no poles in $\Im(g) > 0$], (5) $h(g) \neq \alpha_i$; moreover, $\Im(g) = 0$ is a natural boundary for $h(g)$.

If $\alpha_m = \infty$, we shall use $g = g(x)$; otherwise we put $g = g^*(x)$. $g_1(x)$ is a special case of $g = g(x)$, where the points α_i are the n n-th roots of unity and ∞. The same is true of $g = g_4(x)$, where the points α_i are the n n-th roots of unity, 0 and ∞. $g_2(x)$ and $g_3(x)$ are special cases of $g = g^*(x)$, where the points α_i are the n n-th roots of unity and 0 in the first case, and the n n-th roots of unity in the second.

If we set

$$t(x) = \frac{u^2 g(x) + w}{g(x) + w}, \tag{13}$$

where $w = e^{2\pi i/3}$, then $t(x)$ has properties similar to those of $g(x)$ except that $\Im(g) \geqq 0$ is replaced by $|t| \leqq 1$.

Let $\phi(a_0, a_1, \alpha_1, \alpha_2, \cdots, \alpha_m)$, or more briefly $\phi(a_0, a_1, m)$, be the least possible value of $M(a_0, a_1, \alpha_1, \alpha_2, \cdots, \alpha_m)$ for the extended Landau theorem.

THEOREM 4. *Let a_0 and $a_1 \neq 0$ be two given constants. Let $\phi(a_0, a_1, m) = 0$ if $a_0 = \alpha_i$ $(i = 1, 2, \cdots, m)$; otherwise let*

$$\phi(a_0, a_1, m) = \frac{2\Im\{g(a_0)\}}{|a_1| \cdot |g'(a_0)|} = \frac{1 - |t(a_0)|^2}{|a_1| \cdot |t'(a_0)|}.$$

Every function $F(x)$, regular in the neighborhood of the origin and there defined by $F(x) = a_0 + a_1x + \cdots$, in or on the circle $|x| = \phi(a_0, a_1, m)$ either has a singularity or assumes one of the α_i values. Moreover, the above formula for $\phi(a_0, a_1, m)$ is independent of the branch used for $g(a_0)$.

COROLLARY 1. $\phi(a_0, 1, m) = |a_1| \phi(a_0, a_1, m)$.

If $a_0 = \alpha_i$ $(i = 1, 2, \cdots, m)$, the theorem is obvious, so we shall henceforth suppose that this is not true. The equality in the theorem is readily established by the use of the relation (13). The independence of the branch is established

by the supposition that $K(x)$ and $L(x)$ are two branches of $\dot{g}(x)$. Then they are related by $K = (aL + b)/(cL + d)$, where $ad - bc = 1$, a, b, c, d are real and $\mathfrak{J}(L) \geqq 0$ maps upon $\mathfrak{J}(K) \geqq 0$.

We choose an arbitrary branch of $g(x)$ and form $G(x) = g(y)$, where $y = F(x)$. From the hypothesis on $F(x)$ there exists a positive number η such that for $|x| < \eta$, $F(x)$ is regular and $F(x) \neq \alpha_i$. Hence $G(x)$ is regular and $\mathfrak{J}(G(x)) > 0$ for $|x| < \eta$. Hence

$$G(x) = g(a_0) + a_1 g'(a_0)x + \cdots. \tag{14}$$

We define

$$H(x) = \frac{G(x) - g(a_0)}{G(x) - \bar{g}(a_0)}, \tag{15}$$

where $\bar{g}(a_0)$ is the conjugate imaginary of $g(a_0)$. It is easy to show that $H(x)$ is regular and $|H(x)| < 1$ for $|x| < \eta$. Substituting and expanding, we obtain

$$H(x) = \frac{a_1 g'(a_0)}{2i\,\mathfrak{J}[g(a_0)]}\,x + \cdots. \tag{16}$$

Applying the Cauchy inequality for the first derivative and simplifying, we obtain

$$\eta \leqq \frac{2\,\mathfrak{J}[g(a_0)]}{|a_1|\,|g'(a_0)|}. \tag{17}$$

We observe that in the proof thus far we have made use of only the first three properties of $g(x)$ and hence the result (17) is true for any function possessing those properties.

To complete the proof, we shall exhibit a function $F(x)$ of the required form, which is regular for $|x| < \phi$, and which there stays away from $\alpha_1, \alpha_2, \cdots, \alpha_m$. We do this by defining $H(x)$ by the first term of the series (16) and then $G(x)$ by (15). We use the inverse function of $g = g(x)$ and form

$$F(x) = h[G(x)]. \tag{18}$$

This function $F(x)$ is easily shown to have the required properties and hence the theorem is established. The same theorem is true if we use $g = g(x)$, with the exception that when we form equation (18), $F(x)$ will have simple poles in $|x| = \phi$, and hence the radius in this case is the least possible in a restrictive sense.

Similar theorems are obtainable by the use of $a_0, a_1, \cdots, a_q$ as $q + 1$ given constants.

Corollary 2. *Let $\phi_1(a_0, a_1, n) = 0$ if $a_0^n \neq 1$; otherwise let*

$$\phi_1(a_0, a_1, n) = \frac{1 - |t_1(a_0)|^2}{|a_1| \cdot |t_1'(a_0)|}.$$

Then $F(x)$ in or on the circle $|x| = \phi_1$ either has a singularity or takes on an n-th root of unity value.

COROLLARY 3. *Let* $\phi_2(a_0, a_1, n) = 0$ *if* $a_0 = 0$ *or* $a_0^n = 1$; *otherwise let*

$$\phi_2(a_0, a_1, n) = \frac{1 - |t_2(a_0)|^2}{|a_1| \cdot |t_2'(a_0)|}.$$

Then $F(x)$ *in or on the circle* $|x| = \phi_2$ *either has a singularity, a zero, or assumes an n-th root of unity value.*

COROLLARY 4. *Let* $\phi_3(a_0, a_1, n) = 0$ *if* $a_0^n = 1$; *otherwise let*

$$\phi_3(a_0, a_1, n) = \frac{1 - |t_3(a_0)|^2}{|a_1| \cdot |t_3'(a_0)|}.$$

Then $F(x)$ *in or on the circle* $|x| = \phi_3$ *either has a singularity or assumes an n-th root of unity value.*

We observe that in the third and fourth corollaries the radii are not necessarily the definitive radii, since these two corollaries come under $g = g^*(x)$.

COROLLARY 5. *Let* $\phi_4(a_0, a_1, n) = 0$ *if* $a_0 = 0$ *or* $a_0^n = 1$, *and otherwise let*

$$\phi_4(a_0, a_1, n) = \frac{2\,\mathfrak{J}\{g_4(a_0)\}}{|a_1| \cdot |g_4'(a_0)|}.$$

Then $F(x)$ *in or on the circle* $|x| = \phi_4$ *either has a singularity, a zero, or assumes an n-th root of unity value.*

Moreover, in this last corollary it is easy to show that

$$\phi_4(a_0, a_1, n) = \phi_4(a_0^n, na_0^{n-1}a_1, 1).$$

If we use the identities given in the preceding three sections we obtain the following

THEOREM 5. *Let* $0 < p < 1$. *Then*

$$\lim_{n\to\infty} \phi_1(p, 1, n) = \lim_{n\to\infty} \phi_3(p, 1, n) = 1 - p^2,$$

$$\lim_{n\to\infty} \phi_2(p, 1, n) = \lim_{n\to\infty} \phi_4(p, 1, n) = -2p \log p.$$

Let $1 < p < \infty$. *Then*

$$\lim_{n\to\infty} \phi_1(p, 1, n) = \lim_{n\to\infty} \phi_4(p, 1, n) = 2p \log p,$$

$$\lim_{n\to\infty} \phi_2(p, 1, n) = \lim_{n\to\infty} \phi_3(p, 1, n) = p^2 - 1.$$

Let $\lim_{n\to\infty} \phi_i(a_0, a_1, n) = \psi_i(a_0, a_1)$, $i = 1, 2, 3, 4$. Then in the above theorem we observe that

$$\psi_4(p, 1) \leqq \psi_1(p, 1) \leqq \psi_3(p, 1).$$

THEOREM 6. *Let* $|a_0| \neq 1$. *Then for* $|x| \leqq \psi_1(a_0, a_1)$, $F(x)$ *either has a singularity or a point where* $|F(x)| = 1$.

The following numerical values are for purposes of comparison.

$$\phi_1(0, 1, 3) \approx 2.58, \qquad \phi_3(0, 1, 3) \approx 3.25, \qquad \phi_4(0, 1, 3) = 0,$$

$$\psi_1(0, 1) = 1, \qquad \psi_3(0, 1) = 1, \qquad \psi_4(0, 1) = 0,$$

$$\phi_1(\tfrac{1}{2}, 1, 3) \approx 2.27, \qquad \phi_3(\tfrac{1}{2}, 1, 3) \approx 2.87, \qquad \phi_4(\tfrac{1}{2}, 1, 3) \approx 1.49,$$

$$\psi_1(\tfrac{1}{2}, 1) = 3/4, \qquad \psi_3(\tfrac{1}{2}, 1) = 3/4, \qquad \psi_4(\tfrac{1}{2}, 1) = \log 2.$$

UNIVERSITY OF ILLINOIS.

ON CERTAIN EQUATIONS IN RELATIVE-CYCLIC FIELDS

By Leonard Carlitz

1. Introduction. Let F be a quite arbitrary field—the characteristic may be 0 or some prime p. Let W be a field containing F such that W/F is cyclic of relative degree k. The group of W/F is generated by the substitution S: if α is some quantity in W, we shall use the notation α^S to denote the result of operating on α with S. If then α, β are assigned elements of W, the equations which we shall study are

$$\xi^S = \alpha\xi \tag{1.1}$$

and

$$\eta^S = \alpha\eta + \beta; \tag{1.2}$$

it is of course supposed that ξ and η also are in W.

Suppose $W = F(\vartheta)$, that is, W is generated by adjoining ϑ to F, where ϑ is a root of $f(\vartheta) = 0$, and $f(x)$ is a polynomial with coefficients in F and irreducible in F. It is convenient to assume that the coefficient of the highest power of x in $f(x)$ is unity. Let $\alpha = g(\vartheta)$, where $g(x)$ is a polynomial with coefficients in F. Then we show that (1.1) has a non-trivial solution if and only if

$$R(g, f) = 1;$$

here $R(g, f)$ is the resultant of the polynomials g and f, and may be calculated by means of the division algorithm. If g satisfies certain conditions, a theorem of reciprocity for (1.1) may be stated; in particular, if F is a finite field, this reduces to a known theorem (see §5).

As for equation (1.2), if α is such that (1.1) is not satisfied, then (1.2) has a unique solution. If, however, (1.1) does admit of a non-trivial solution, then we may assume $\alpha = 1$, and our equation becomes

$$\xi^S = \xi + \beta. \tag{1.3}$$

If now we put $\beta = h(\vartheta)$, where $h(x)$ is a properly chosen polynomial in F, then we prove that (1.3) is solvable if and only if the coefficient of x^{k-1} in $h(x)\,f'(x)$, reduced modulo $f(x)$, is zero; here $f'(x)$ denotes the derivative of $f(x)$.

In §4 some properties of the solutions of (1.3) are derived. Finally in §5 we assume F to be a finite field and the result for (1.3) as well as for (1.1) is seen to reduce to a known theorem.

Received October 6, 1936; presented to the American Mathematical Society, April 11, 1936.

2. **The equation** (1.1). For α^S as defined above, we note that

$$(\alpha\beta)^S = \alpha^S\beta^S, \qquad (\alpha + \beta)^S = \alpha^S + \beta^S,$$

where α, β are arbitrary elements of W. For λ in F, $\lambda^S = \lambda$; conversely, if the substitution S leaves some element λ of W unchanged, then λ must be in F. For arbitrary α in W we may say only

$$\alpha^{S^k} = \alpha. \tag{2.1}$$

We may evidently assume in (1.1) that $\alpha \neq 0$. We now define the following function of α:

$$\chi = \chi(\alpha) = \alpha\alpha^S \cdots \alpha^{S^{k-1}} = \alpha^{1+S+\cdots+S^{k-1}}. \tag{2.2}$$

Thus it is evident that

$$\chi^S = \alpha^{S+S^2+\cdots+S^k} = \alpha^{1+S+\cdots+S^{k-1}},$$

by (2.1), so that $\chi^S = \chi$, and therefore χ is in F. Clearly if $\alpha \neq 0$, then $\chi \neq 0$. Note also that

$$\chi(\alpha\beta) = \chi(\alpha)\chi(\beta),$$

$$\chi(1) = 1, \qquad \chi(\lambda) = \lambda^k$$

for α, β in W, λ in F. If now we assume that (1.1) has a non-trivial solution ($\xi \neq 0$), then

$$\xi^{S^2} = \alpha^S\xi^S = \alpha^{1+S}\xi,$$

$$\xi^{S^3} = \alpha^{S^2}\xi^{S^2} = \alpha^{1+S+S^2}\xi,$$

$$\xi^{S^k} = \xi = \alpha^{1+S+\cdots+S^{k-1}}\xi = \chi\xi,$$

so that $\chi = 1$; that is, a necessary condition that (1.1) be solvable in W is $\chi(\alpha) = 1$.

To show that this condition is also sufficient, consider for arbitrary β the sum

$$A = \sum_{j=0}^{k-1} \beta^{S^j}\alpha^{-(1+S+\cdots+S^{j-1})}. \tag{2.3}$$

Then applying S:

$$\begin{aligned}
A^S &= \sum_0^{k-1} \beta^{S^{j+1}}\alpha^{-(S+S^2+\cdots+S^j)} \\
&= \alpha \sum_0^{k-1} \beta^{S^{j+1}}\alpha^{-(1+S+\cdots+S^j)} \\
&= \alpha \sum_1^{k} \beta^{S^j}\alpha^{-(1+S+\cdots+S^{j-1})} \\
&= \alpha \sum_0^{k-1} \beta^{S^j}\alpha^{-(1+S+\cdots+S^{j-1})} + \beta\left(\frac{1}{\chi(\alpha)} - 1\right),
\end{aligned}$$

that is,

$$A^s - \alpha A = \beta\left(\frac{1}{\chi(\alpha)} - 1\right). \tag{2.4}$$

Hence for $\chi(\alpha) = 1$, A satisfies our equation (1.1). It remains to show that for properly chosen β, $A \neq 0$. Now if $A = 0$ for all β, in particular it vanishes for all

$$\beta^i \qquad (i = 0, \cdots, k-1);$$

substitution in (2.3) leads to

$$\sum_{j=0}^{k-1} (\beta^i)^{s^j} \alpha^{-(1+s+\cdots+s^{j-1})} = 0 \qquad (i = 0, \cdots, k-1).$$

In other words, the set of linear equations

$$\sum_{j=0}^{k-1} (\beta^i)^{s^j} y^j = 0 \qquad (i = 0, \cdots, k-1)$$

has a non-trivial solution, and therefore the determinant

$$|(\beta^i)^{s^j}| = 0 \qquad (i, j = 0, \cdots, k-1);$$

that is to say, the relative discriminant of β vanishes. Thus by properly choosing β, A as defined by (2.3) is different from zero. We have therefore proved[1]

THEOREM 2.1. *A necessary and sufficient condition that* (1.1) *have a solution other than* $\xi = 0$ *is furnished by* $\chi(\alpha) = 1$, *where* $\chi(\alpha)$ *is defined by* (2.1).

We may now derive the criterion mentioned in the Introduction. For f and g as defined in §1, we have

$$\alpha = g(\vartheta), \quad \alpha^s = g(\vartheta^s), \quad \cdots,$$

so that

$$\begin{aligned} \chi(\alpha) &= \alpha^{1+s+\cdots+s^{k-1}} \\ &= g(\vartheta)g(\vartheta^s)\cdots g(\vartheta^{s^{k-1}}). \end{aligned} \tag{2.5}$$

Now assume the factorization

$$g(\chi) = \lambda \prod_{\omega} (\chi - \omega), \qquad \lambda \text{ in } F,$$

in some field W_1 containing W. Then by (2.5) we have the factorization

$$\begin{aligned} \chi(\alpha) &= \lambda^k \Pi\, \Pi(\vartheta - \omega) \\ &= R(g(x), f(x)), \end{aligned} \tag{2.6}$$

where $R(g, f)$ is the resultant of g and f. Applying Theorem (2.1), we now have

[1] Cf. D. Hilbert, *Jahresbericht der Deutschen Mathematiker Vereinigung*, vol. 4 (1894–95), pp. 271–272.

THEOREM 2.2. *Equation* (1.1) *has a solution in* W *other than* $\xi = 0$ *if and only if the resultant of* g *and* f *is unity*:

$$R(g, f) = 1.$$

If $\xi \neq 0$ *is a particular solution of* (1.1), *then the general solution is* $\lambda\xi$, λ *arbitrary in* F.

We may suppose $R(g, f)$ defined by (2.5) and (2.6). It is assumed that f is *primary*, that is, the coefficient of the highest power of x is unity. More generally for $f = \Pi(x - \vartheta)$, we write

$$R(g, f) = \prod_{\vartheta} g(\vartheta). \tag{2.7}$$

From (2.7) it is clear that $g \equiv h \pmod f$ implies $R(g, f) = R(h, f)$. Also for λ "constant", $R(\lambda, f) = \lambda^k$ by (2.7), where k is the degree of f. For arbitrary g, h we have $R(gh, f) = R(g, f)R(h, f)$. Thus for f, g primary of degree k, l respectively, it follows that

$$R(g, f) = (-1)^{kl}R(f, g). \tag{2.8}$$

It is now clear how $R(g, f)$ may be calculated by means of the division algorithm, and thus by (2.6) χ may be evaluated. The question arises whether (2.8) may be interpreted as a reciprocity relation for χ. Apparently this cannot be done in general. If, however, the roots of g (assumed irreducible in F) adjoined to F generate a cyclic super-field of F, then clearly $R(f, g)$ has the same interpretation as $R(g, f)$ and a reciprocity theorem may be stated for χ. (Cf. §5.)

3. **The equation** (1.2). Assume first that $\chi(\alpha) \neq 1$; then (2.3) implies

$$\left(\frac{\chi(\alpha)\, A}{\chi(\alpha) - 1}\right)^s = \alpha\, \frac{\chi(\alpha)\, A}{\chi(\alpha) - 1} + \beta,$$

so that in this case (1.2) has the solution

$$\eta = \frac{\chi(\alpha)}{\chi(\alpha) - 1}\, A.$$

Further, this is the only solution, for

$$\eta_1^s = \alpha\eta_1 + \beta, \qquad \eta_2^s = \alpha\eta_2 + \beta$$

leads to

$$(\eta_1 - \eta_2)^s = \alpha(\eta_1 - \eta_2),$$

and since $\chi(\alpha) \neq 1$, Theorem 2.1 implies $\eta_1 = \eta_2$.

We may therefore suppose that $\chi(\alpha) = 1$; then there exists a $\gamma \neq 0$ such that

$$\alpha = \gamma^{s-1}.$$

In (1.2) put $\eta = \gamma\zeta$: the equation becomes

$$\zeta^s = \zeta + \beta\gamma^{-s}.$$

We therefore consider in the following the equation (1.3):

$$\eta^S = \eta + \beta.$$

In place of $\chi(\alpha)$ we now define the quantity $\rho = \rho(\beta)$ by

$$\rho(\beta) = \beta + \beta^s + \cdots + \beta^{s^{k-1}}; \tag{3.1}$$

then

$$\rho^s = \beta^s + \rho^{s^2} + \cdots + \beta^{s^k} = \rho,$$

and $\rho(\beta)$ lies in F. Evidently $\rho(\beta)$ has the properties

$$\begin{gathered} \rho(\beta + \gamma) = \rho(\beta) + \rho(\gamma), \\ \rho(\lambda\beta) = \lambda\rho(\beta), \qquad \rho(\lambda) = k\lambda, \end{gathered} \tag{3.2}$$

for λ a quantity in F.

From (3.1) it is easily seen that a necessary condition that (1.3) have a solution is $\rho(\beta) = 0$. To show that this condition is also sufficient, consider the sum

$$B = \sum_{j=0}^{k-1} \alpha^{s^j} \beta_j,$$

where

$$\beta_j = \beta + \beta^s + \cdots + \beta^{s^j}, \qquad \beta_0 = \beta. \tag{3.3}$$

Then applying S,

$$\begin{aligned} B^S &= \sum_0^{k-1} \alpha^{s^{j+1}} (\beta_{j+1} - \beta) \\ &= \sum_0^{k-1} \alpha^{s^j} (\beta_j - \beta) + \alpha^{s^k} (\beta_k - \beta) \\ &= B - \beta\rho(\alpha) + \alpha\rho(\beta), \end{aligned}$$

for by (3.3),

$$\beta_k - \beta = \beta^s + \cdots + \beta^{s^k} = \rho^s(\beta) = \rho(\beta).$$

Hence the identity for arbitrary α,

$$B^S - B = \alpha\rho(\beta) - \beta\rho(\alpha). \tag{3.4}$$

Let us now assume $\rho(\beta) = 0$; then we have

$$B^S - B = \beta\rho(-\alpha).$$

If then we can find an α such that $\rho(-\alpha) = 1$, B will satisfy (1.3). We shall prove slightly more:

THEOREM 3.1. *For arbitrary λ in F, there exists a γ in W such that*

$$\rho(\gamma) = \lambda.$$

Assume that $\rho(\gamma) = 0$ for all γ in W. Then in particular for the quantities

$$\gamma^i \qquad (i = 0, \cdots, k-1),$$

we have $\rho(\gamma^i) = 0$, that is,

$$\sum_{j=0}^{k-1} \gamma^{is^j} = 0 \qquad (i = 0, \cdots, k-1),$$

whence exactly as in §2, the relative discriminant

$$| \gamma^{is^j} | = 0 \qquad (i, j = 0, \cdots, k-1).$$

Therefore for some γ in W, $\rho(\gamma) = \mu \neq 0$. By (3.2),

$$\rho\left(\frac{1}{\mu}\gamma\right) = 1, \qquad \rho\left(\frac{\lambda}{\mu}\gamma\right) = \lambda,$$

which proves the theorem. As we have already noticed, this implies the following

THEOREM 3.2. *A necessary and sufficient condition for the solvability of* (1.3) *is furnished by $\rho(\beta) = 0$, where $\rho(\beta)$ is defined by* (3.1).

To derive the criterion stated in the Introduction, we need the following easily proved identity:

$$\frac{f'(x)}{f(x)} = \sum_{j=0}^{k-1} \frac{1}{x - \vartheta^{s^j}}. \tag{3.5}$$

Since

$$\frac{1}{x - \vartheta} = \frac{1}{x} + \frac{\vartheta}{x^2} + \cdots + \frac{\vartheta^{m-1}}{x^m} + \frac{\vartheta^m}{x^m}\frac{1}{x - \vartheta},$$

we have from (3.5)

$$\frac{f'(x)}{f(x)} = \frac{1}{x}\rho(1) + \frac{1}{x^2}\rho(\vartheta) + \cdots + \frac{1}{x^m}\rho(\vartheta^{m-1}) + \frac{1}{x^m}\sum_j \frac{\vartheta^{ms^j}}{x - \vartheta^{s^j}},$$

and therefore

$$\begin{aligned} x^m f'(x) &\equiv \sum_j \vartheta^{ms^j} \frac{f(x)}{x - \vartheta^{s^j}} && (\operatorname{mod} f(x)) \\ &\equiv \sum_j \vartheta^{ms^j} (x^{k-1} + \cdots) && (\operatorname{mod} f(x)), \end{aligned}$$

the dots indicating powers of x of exponent $< k - 1$. Comparing coefficients of x^{k-1} on both sides of this congruence, we see that $\rho(\vartheta^m)$ is the coefficient of x^{k-1} in the product $x^m f'(x)$ reduced modulo $f(x)$. If, then, $h(x)$ is a polynomial

with coefficients in F, it follows upon application of (3.2) that $\rho(h(\vartheta))$ is the coefficient of x^{k-1} in the residue of $h(x)f'(x) \pmod{f(x)}$. Thus we have

THEOREM 3.3. *If $\beta = h(\vartheta)$, where $h(x)$ is a polynomial in F, then $\rho(\beta)$ is the coefficient of x^{k-1} in $h(x)f'(x)$ reduced modulo $f(x)$. In particular,* (1.3) *is solvable in W if and only if that coefficient is zero. If η is a particular solution of* (1.3), *then the general solution is $\eta + \lambda$, λ arbitrary in F.*

4. **Some properties of the sum** B. We shall use the fuller notation (α, β) in place of B:

$$(\alpha, \beta) = \sum_{0}^{k-1} \alpha^{s^i} \beta_i, \tag{4.1}$$

where β_i is defined by (3.3). It is evident that

$$\begin{aligned} (\alpha + \beta, \gamma) &= (\alpha, \gamma) + (\beta, \gamma), \\ (\alpha, \beta + \gamma) &= (\alpha, \beta) + (\alpha, \gamma), \\ (\lambda\alpha, \beta) &= \lambda(\alpha, \beta) = (\alpha, \lambda\beta), \end{aligned} \tag{4.2}$$

for λ in F. We derive a formula connecting (α, β) and (β, α). If in (3.4) we interchange α and β and add, we have

$$(\alpha, \beta)^s + (\beta, \alpha)^s = (\alpha, \beta) + (\beta, \alpha),$$

so that the sum $(\alpha, \beta) + (\beta, \alpha)$ is in F. To get an explicit expression we proceed as follows. The product

$$\begin{aligned} \rho(\alpha)\,\rho(\beta) &= \sum_{i,j=0}^{k-1} \alpha^{s^i} \beta^{s^j} \\ &= \sum_{i \leqq j} + \sum_{i \geqq j} - \sum_{i=j}. \end{aligned} \tag{4.3}$$

But

$$\sum_{i \leqq j} \alpha^{s^i} \beta^{s^j} = \sum_j \beta^{s^j} \sum_{i \leqq j} \alpha^{s^i} = (\beta, \alpha), \tag{4.4}$$

by (4.1) and (3.3). Similarly

$$\sum_{i \geqq j} \alpha^{s^i} \beta^{s^j} = \sum_i \alpha^{s^i} \sum_{j \leqq i} \beta^{s^j} = (\alpha, \beta), \tag{4.5}$$

while

$$\sum_{i=j} \alpha^{s^i} \beta^{s^j} = \sum_j (\alpha\beta)^{s^j} = \rho(\alpha\beta). \tag{4.6}$$

Combining (4.3), $\cdots$, (4.6), we have at once

$$(\alpha, \beta) + (\beta, \alpha) = \rho(\alpha\beta) + \rho(\alpha)\rho(\beta), \tag{4.7}$$

which indicates explicitly that the sum in the left member is in F.

Again from (3.4),

$$(\alpha, \beta)^s - (\alpha, \beta) = \alpha\rho(\beta) - \beta\rho(\alpha), \tag{4.8}$$

for arbitrary α, β. In particular for $\rho(\beta) = 0$,

$$(\alpha, \beta)^s - (\alpha, \beta) = -\beta\rho(\alpha).$$

Indeed, for $\beta = \gamma^s - \gamma$, by (3.3),

$$(\alpha, \beta) = \Sigma\alpha^{s^j}(\gamma^{s^{j+1}} - \gamma) = \rho(\alpha\gamma^s) - \gamma\rho(\alpha);$$

that is,

$$(\alpha, \gamma^s - \gamma) = \rho(\alpha\gamma^s) - \gamma\rho(\alpha). \tag{4.9}$$

If also $\rho(\alpha) = 0$, then

$$(\alpha, \beta) = \rho(\alpha\gamma^s). \tag{4.10}$$

To determine (α, β) for $\rho(\alpha) = 0$, β arbitrary, we apply (4.7) and (4.9). For $\rho(\alpha) = 0$, (4.7) becomes

$$(\alpha, \beta) = \rho(\alpha\beta) - (\beta, \alpha).$$

Put $\alpha = \gamma^s - \gamma$; then by (4.9)

$$(\beta, \alpha) = \rho(\beta\gamma^s) - \gamma\rho(\beta),$$

so that

$$(\alpha, \beta) = -\rho(\beta\gamma) + \gamma\rho(\beta). \tag{4.11}$$

If in (4.10) we put $\alpha = \beta$, we have

$$(\beta, \beta) = \rho(\beta\gamma^s), \tag{4.12}$$

while from (4.11),

$$(\beta, \beta) = -\rho(\beta\gamma). \tag{4.13}$$

However,

$$\begin{aligned} \rho(\beta\gamma^s) &= \rho((\gamma^s - \gamma)\gamma^s) \\ &= \rho(\gamma^{2s} - \gamma^{1+s}) \\ &= \rho(\gamma^2 - \gamma^{1+s}) \\ &= \rho((\gamma - \gamma^s)\gamma) = -\rho(\beta\gamma), \end{aligned}$$

so that (4.12) and (4.13) are identical.

We may now evaluate (β, β) for arbitrary β. Since the case $\rho(\beta) = 0$ has been disposed of, we assume $\rho(\beta) \neq 0$; in particular, let $\rho(\beta) = 1$. Suppose also that $\rho(\delta) = 1$. Then by (4.8),

$$(\beta, \delta)^s - (\beta, \delta) = \beta - \delta;$$

put $\gamma = (\beta, \delta)$ so that $\beta - \delta = \gamma^S - \gamma$. We have on the one hand

$$(\beta - \delta, \delta) = (\gamma^S - \gamma, \delta) = \gamma + \rho(\gamma\delta).$$

But on the other hand,

$$(\beta - \delta, \delta) = (\beta, \delta) - (\delta, \delta) = \gamma - (\delta, \delta).$$

Comparing the two expressions for $(\beta - \delta, \delta)$, we see that

$$(\delta, \delta) = -\rho(\gamma\delta), \tag{4.14}$$

where $\gamma = (\beta, \delta)$ and β is arbitrary except for $\rho(\beta) = 1$.

Finally we consider the expression

$$\Gamma = \rho(\alpha)\cdot(\beta, \gamma) + \rho(\beta)\cdot(\gamma, \alpha) + \rho(\gamma)\cdot(\alpha, \beta). \tag{4.15}$$

By (4.7),

$$\begin{aligned}&\rho(\beta)(\gamma, \alpha) + \rho(\gamma)(\alpha, \beta)\\ &\qquad = \rho(\alpha)\rho(\beta)\rho(\gamma) + \rho(\alpha\beta)\rho(\gamma) + \rho(\beta)(\gamma, \alpha) - (\beta, \alpha)\rho(\gamma);\end{aligned} \tag{4.16}$$

but

$$\begin{aligned}\rho(\beta)(\gamma, \alpha) - \rho(\gamma)(\beta, \alpha) &= (\gamma\rho(\beta) - \beta\rho(\gamma), \alpha)\\ &= -((\beta, \gamma)^S - (\beta, \gamma), \alpha) \qquad \text{by (4.8)},\\ &= \rho\{\alpha(\beta, \gamma)\} - \rho(\alpha)(\beta, \gamma).\end{aligned} \tag{4.17}$$

Substituting from (4.17) in (4.16), we have

$$\begin{aligned}&\rho(\alpha)(\beta, \gamma) + \rho(\beta)(\gamma, \alpha) + \rho(\gamma)(\alpha, \beta)\\ &\qquad = \rho(\alpha)\rho(\beta)\rho(\gamma) + \rho(\alpha\beta)\rho(\gamma) + \rho\{\alpha(\beta, \gamma)\},\end{aligned} \tag{4.18}$$

so that in particular Γ as defined by (4.15) is in F.

5. **Application to finite field** F. Let F be a Galois field $GF(p^n)$ of order p^n, where p is an arbitrary prime. Then for $f(x)$ an irreducible polynomial in $GF(p^n)$, a cyclic extension W of $GF(p^n)$ is generated by adjoining a root ϑ of $f(\vartheta) = 0$ to the Galois field. We remark that W may also be defined as the field formed by the complete set of residues modulo $f(x)$. For the present case it is convenient to use the latter interpretation so that we shall speak of congruences (mod $f(x)$) rather than equations in W. Note that for $F = GF(p^n)$, W is "absolute" cyclic, that is, W is cyclic relative to $GF(p)$ as well as relative to $GF(p^n)$. Clearly W is also a finite field. The substitution S that generates the cyclic group of W/F may be identified with the operation of taking the p^n-th power:

$$\alpha^S \to \alpha^{p^n};$$

however, other interpretations are possible. For example, if as above the relative degree of W/F is k, we may take

$$\alpha^S \to \alpha^{p^{n(k-1)}},$$

or generally

$$\alpha^S \to \alpha^{p^{nr}},$$

where r is prime to k.

For brevity we limit ourselves to the first definition of S; then our equations (1.1) and (1.3) become

$$X^{p^n} \equiv gX \pmod{f} \tag{5.1}$$

and

$$X^{p^n} \equiv X + h \pmod{f}, \tag{5.2}$$

respectively, where g, h, X, Y are polynomials with coefficients in $GF(p^n)$.

For the equation (5.1) it is customary to use the notation $\{g/f\}$ in place of χ, so that we have

$$\left\{\frac{g}{f}\right\} \equiv g^{1+p^n+\cdots+p^{n(k-1)}} \pmod{f}. \tag{5.3}$$

Then by the proof of Theorem 2.2,

$$\left\{\frac{g}{f}\right\} \equiv R(g, f).$$

By the remark at the end of §2, a reciprocity theorem for $\{g/f\}$ may be stated. Indeed, for g irreducible in $GF(p^n)$, the set of residues (mod g) form a field that is cyclic relative to $GF(p^n)$. Hence we have the following[2]

THEOREM 5.1. *If f and g are primary irreducible polynomials in $GF(p^n)$ of degree k and l, respectively, then*

$$\left\{\frac{f}{g}\right\} = (-1)^{kl}\left\{\frac{g}{f}\right\},$$

where $\{g/f\}$ is defined by (5.3).

For the equation (5.2) we have

$$\rho(h) \equiv h + h^{p^n} + \cdots + h^{p^{n(k-1)}} \pmod{f}; \tag{5.4}$$

then Theorem 3.3 implies[3]

THEOREM 5.2. *$\rho(h)$, as defined by* (5.4), *is equal to the coefficient of x^{k-1} in the product hf' reduced* (mod f). *The equation* (5.2) *is solvable—in polynomials with coefficients in $GF(p^n)$—if and only if that coefficient vanishes.*

INSTITUTE FOR ADVANCED STUDY AND DUKE UNIVERSITY.

[2] F. K. Schmidt, *Erlanger Sitzungsberichte*, vols. 58–59 (1928), pp. 159–172.

[3] L. Carlitz, this journal, vol. 1 (1935), pp. 164–168; also Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 844–846.

ON FACTORABLE POLYNOMIALS IN SEVERAL INDETERMINATES

BY LEONARD CARLITZ

1. **Introduction.** In this paper we consider a class of polynomials in several indeterminates with coefficients in a Galois field $GF(p^n)$, such that each polynomial may be completely factored into a product of linear factors in *some* Galois field $GF(p^{nn'})$, say. For the case of a single indeterminate a body of theorems[1] exists, and the purpose of this paper is to extend these theorems, whenever possible, to the case of several indeterminates. As will be seen in several cases, certain theorems are capable of extension, but the proof for the case of a single indeterminate is no longer applicable, and new methods become necessary. This is true in particular of the formula for the product of all (factorable) polynomials of fixed degree. Again, in the case of a single indeterminate, the form of a polynomial is known explicitly; in the case of several indeterminates, the definition is in terms of an intrinsic property, and thus it seems necessary to deal first with irreducible polynomials and from them go on to arbitrary polynomials.

In the case of polynomials in a single indeterminate x, as is well known, the quantity

$$x^{p^{ns}} - x \tag{1.1}$$

is fundamental.[2] In the extended case this is replaced by a certain determinant. Thus for example, for two and three indeterminates, we have

$$\begin{vmatrix} 1 & x & y \\ 1 & x^{p^{ns}} & y^{p^{ns}} \\ 1 & x^{p^{2ns}} & y^{p^{2ns}} \end{vmatrix}, \qquad \begin{vmatrix} 1 & x & y & z \\ 1 & x^{p^{ns}} & y^{p^{ns}} & z^{p^{ns}} \\ 1 & x^{p^{2ns}} & y^{p^{2ns}} & z^{p^{2ns}} \\ 1 & x^{p^{3ns}} & y^{p^{3ns}} & z^{p^{3ns}} \end{vmatrix}, \tag{1.2}$$

respectively. Certain formulas in the case of a single indeterminate carry over to the extended case by merely substituting the proper expression (1.2) for (1.1). In particular this is true for the product of irreducible polynomials and the product of all polynomials of fixed degree.

The number of (primary) irreducible polynomials of degree s in a single indeterminate, with coefficients in the $GF(p^n)$, is determined by the familiar expression

$$\psi(s, p^n) = \sum_{s=d\delta} \mu(\delta)\, p^{nd},$$

Received October 12, 1936.

[1] For the classic theorems, see L. E. Dickson, *Linear Groups*, 1901, pp. 3–54.

[2] Dickson, loc. cit.

where $\mu(\delta)$ is the Möbius μ-function. Now the number of irreducible polynomials in x, y (y actually appearing) will be shown to be $\psi(s, p^{2n})$; for k indeterminates (a fixed one appearing in each polynomial) the number is $\psi(s, p^{kn})$. Similarly, the number of primary polynomials of degree m in a single indeterminate is p^{nm}; for k indeterminates the number of factorable polynomials (in which x_k^m actually occurs) is p^{nmk}.

2. **Some definitions.** According to the definition above, a polynomial $M = M(x_1, \cdots, x_k)$ is factorable provided it can be written in the form

$$M(x_1, \cdots, x_k) = \prod_{i=1}^{m} (\alpha_{i0} + \alpha_{i1} x_1 + \cdots + \alpha_{ik} x_k), \tag{2.1}$$

where the α_{ij} lie in *some* Galois field, $GF(p^{nn'})$, say. The coefficients of M are all assumed to be in a fixed $GF(p^n)$. The degree of M is evidently m: we write $m = \deg M$. If x_k^m actually appears in M, then M is called *primary* provided the coefficient of x_k^m is unity. More generally, let us write in place of (2.1)

$$M = \prod_{i=1}^{m} (\alpha_{i0} + \cdots + \alpha_{ik_i} x_{k_i}), \qquad \prod_{i=1}^{m} \alpha_{ik_i} \neq 0;$$

then M is primary if and only if the product $\Pi\alpha_{ik_i} = 1$. From the definition it follows readily that the product of two primary polynomials is itself primary.[3]

It is occasionally convenient to introduce an additional indeterminate x_0. Our polynomials then become homogeneous; (2.1) now becomes

$$M(x_0, x_1, \cdots, x_k) = \prod(\alpha_{i0}x_0 + \alpha_{i1}x_1 + \cdots + \alpha_{ik}x_k).$$

However, in counting the number of polynomials of a certain set, or in reckoning the degree of a polynomial, we shall always suppose $x_0 \equiv 1$; thus $x_0x_1x_2$ is of degree two.

In the second place it is convenient to distinguish those polynomials of degree m that actually contain the term x_k^m. We shall generally denote such a polynomial by M^*. Similarly if $g(m)$ denotes the number of polynomials M of degree m that have a certain property, then $g^*(m)$ will denote the number of M^* of degree m having the property in question.

3. **Irreducible polynomials.** Assume $P = P(x_1, \cdots, x_k)$, a factorable polynomial of degree s with coefficients in $GF(p^n)$ and irreducible in $GF(p^n)$. Then by (2.1)

$$P = (\alpha_0 + \alpha_1x_1 + \cdots + \alpha_kx_k)A, \tag{3.1}$$

[3] An equivalent definition is the following: M is primary if and only if the coefficient of the monomial of highest rank occurring in M is unity. A monomial $x_1^{i_1} \cdots x_k^{i_k}$ is of higher rank than $x_1^{j_1} \cdots x_k^{j_k}$ provided the first non-vanishing difference in the sequence $i_k - j_k, \cdots, i_1 - j_1$, is positive.

where α_j are in some $GF(p^{nn'})$, and A is a polynomial with coefficients in $GF(p^{nn'})$. If we replace each coefficient in (3.1) by its p^n-th power, P remains unchanged, from which it follows that

$$P = (\alpha_0^{p^n} + \alpha_1^{p^n} x_1 + \cdots + \alpha_k^{p^n} x_k)A',$$

so that P is divisible by $\alpha_0^{p^n} + \cdots + \alpha_k^{p^n} x_k$. Similarly P is divisible by each of

$$\alpha_0^{p^{nj}} + \alpha_1^{p^{nj}} x_1 + \cdots + \alpha_k^{p^{nj}} x_k \qquad (j = 0, 1, \cdots). \tag{3.2}$$

Now there is only a finite number of distinct linear forms (3.2); let the number be t. Clearly the product

$$G = \prod_{j=0}^{t-1} (\alpha_0^{p^{nj}} + \cdots + \alpha_k^{p^{nj}} x_k)$$

is in $GF(p^n)$. On the other hand, P is evidently a multiple of G. Thus $P = G$ and $s = t$. Hence we have the factorization

$$P = \prod_{j=0}^{s-1} (\alpha_0^{p^{nj}} + \cdots + \alpha_k^{p^{nj}} x_k). \tag{3.3}$$

Note therefore that if P contains x_k at all, it must contain x_k^s, so that by the definition at the end of §2, $P = P^*$.

We now determine s in terms of the α_j appearing in (3.3). If α is contained in $GF(p^{nf})$ but no $GF(p^{ne})$ for $1 \leqq e < f$, we shall call f the degree of α relative to $GF(p^n)$: $f = \deg \alpha$. Then for $f_j = \deg \alpha_j$ $(j = 0, \cdots, k)$, we prove that

$$s = [f_0, f_1, \cdots, f_k], \tag{3.4}$$

the least common multiple of $f_0, f_1, \cdots, f_k$. Indeed this follows immediately from the fact that s is the number of distinct forms (3.2), that is, s is the smallest positive t such that

$$\alpha_j^{p^{nt}} = \alpha_j \qquad (j = 0, \cdots, k). \tag{3.5}$$

Since for fixed j, f_j is the smallest positive t for which (3.5) holds, it is evident that s is the least common multiple of $f_0, f_1, \cdots, f_k$.

If now we start with $k + 1$ quantities $\alpha_0, \alpha_1, \cdots, \alpha_k$ of degree $f_0, f_1, \cdots, f_k$, respectively; define s by (3.4), and form the product in the right member of (3.3), the polynomial thus formed is in $GF(p^n)$, and further is irreducible in $GF(p^n)$. For if we assume that it equals AB, A and B polynomials in $GF(p^n)$, we may show that either A or B coincides with P: Assume A divisible by $\alpha_0 + \cdots + \alpha_k x_k$; then by the argument used at the beginning of this section, A is divisible by each of the linear forms (3.2), and therefore A is identical with P. Thus we have proved the following

Theorem 3.1. *A (factorable) polynomial P of degree s is irreducible if and only if it satisfies* (3.3) *and s satisfies* (3.4).

In order to determine the product of *all* irreducible polynomials of fixed degree,

we make use of a formula due to E. H. Moore. It is convenient to bring x_0 in at this point. In the (homogeneous) linear form

$$\lambda = \alpha_0 x_0 + \alpha_1 x_1 + \cdots + \alpha_k x_k,$$

where all the α_i are in $GF(p^{ns})$, assume the non-vanishing α_j of *greatest* subscript is equal to unity; that is, assume that one of the following mutually exclusive cases obtains:

$$\begin{aligned}
&\alpha_k = 1;\\
&\alpha_k = 0; \qquad \alpha_{k-1} = 1;\\
&\alpha_k = \alpha_{k-1} = 0, \qquad \alpha_{k-2} = 1;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&\alpha_k = \cdots = \alpha_1 = 0, \qquad \alpha_0 = 1.
\end{aligned}$$

Then the product of all λ satisfying one of these conditions is

$$\prod \lambda = D^s(x_0 x_1 \cdots x_k) = | x_i^{p^{nsj}} | \qquad (i, j = 0, \cdots, k), \tag{3.6}$$

by the formula referred to. On the other hand, in the product of the linear forms λ group together all forms for which the least common multiple of $f_0, f_1, \cdots, f_k = t$, some fixed divisor of s. By Theorem 3.1, it is evident that the product of these forms is identical with the product of all irreducible polynomials of degree t: this product will be denoted by $\Theta(t)$. Comparison with (3.6) leads to the fundamental relation

$$D^s(x_0 x_1 \cdots x_k) = \prod_{t \mid s} \Theta(t). \tag{3.7}$$

From this $\Theta(s)$ is determined by means of the well-known inversion formula:

$$\Theta(s) = \prod_{s = ef} \{D^e(x_0 x_1 \cdots x_k)\}^{\mu(f)}. \tag{3.8}$$

As for $\Theta^*(s)$, the product of the irreducible polynomials of degree s actually containing x_k (and therefore, by the remark following (3.3), necessarily containing x_k^s), we have the formulas

$$\frac{D^s(x_0 x_1 \cdots x_k)}{D^s(x_0 \cdots x_{k-1})} = \prod_{t \mid s} \Theta^*(t) \tag{3.9}$$

and

$$\Theta^*(s) = \prod_{s = ef} \left\{\frac{D^e(x_0 x_1 \cdots x_k)}{D^e(x_0 \cdots x_{k-1})}\right\}^{\mu(f)}. \tag{3.10}$$

This proves the

THEOREM 3.2. *If $\Theta(s) = \Theta(s; x_0 x_1 \cdots x_k)$ denotes the product of the primary irreducible (factorable) polynomials of degree s, and $\Theta^*(s) = \Theta^*(s; x_0 x_1 \cdots x_k)$*

the product of those containing x_k, *then* $\Theta(s)$ *and* $\Theta^*(s)$ *are determined by* (3.8) *and* (3.10), *respectively. Further,*

$$\Theta^*(s; x_0x_1 \cdots x_k) = \frac{\Theta(s; x_0x_1 \cdots x_k)}{\Theta(s; x_0 \cdots x_{k-1})}. \tag{3.11}$$

To determine $\psi_k(s, p^n)$, the number of irreducible P of degree s, it is only necessary to compare the degree of the two members of (3.8). Thus we have immediately

$$\psi_k(s, p^n) = \frac{1}{s} \sum_{s=ef} \mu(e)\,(p^{nkf} + p^{nk(f-1)} + \cdots + p^{nk}), \tag{3.12}$$

and by (3.10) or (3.11),

$$\psi_k^*(s, p^n) = \frac{1}{s} \sum_{s=ef} \mu(e)\, p^{nkf}. \tag{3.13}$$

Therefore it follows that

$$\psi_k^*(s, p^n) = \psi_1^*(s, p^{nk}) = \psi(s, p^{nk}), \tag{3.14}$$

the well known expression for the number of irreducible P in a *single* indeterminate, *but with coefficients in* $GF(p^{nk})$. We may state

THEOREM 3.3. *The number of irreducible polynomials is determined by* (3.12); *the number of those containing* x_k *by* (3.13); *the latter is identical with the number of irreducible polynomials in a single indeterminate, with coefficients from the larger field* $GF(p^{nk})$.

4. **An identity for ψ.** We now count the irreducible P in a different way. For simplicity suppose $k = 2$: the reasoning is quite general and applies without change for arbitrary k. We consider those $P = P^*(x, y)$ actually containing y; then by Theorem 3.1,

$$P = \prod_{j=0}^{s-1} (\alpha^{p^{nj}} + \beta^{p^{nj}} x + y), \tag{4.1}$$

where

$$e = \deg \alpha, \qquad f = \deg \beta, \qquad s = [e, f].$$

We next determine, in the $GF(p^{ns})$, the number of α of degree e relative to $GF(p^n)$. Since each such α gives rise to a polynomial[4]

$$Q = Q(y) = \prod_{j=0}^{e-1} (\alpha^{p^{nj}} + y),$$

irreducible in $GF(p^n)$, it follows that the number of such α is e times the number of Q; but the number of Q of degree e is $\psi(e, p^n)$, and therefore the number of α in $GF(p^{ns})$ of degree e relative to $GF(p^n)$

[4] Dickson, loc. cit., p. 21.

$$= e\,\psi\,(e, p^n) = \sum_{e=d\,\delta} \mu(d)\, p^{n\delta}. \tag{4.2}$$

Similarly the number of $\beta = f\psi(f, p^n)$. Finally, since the s pairs

$$\alpha^{p^{nj}}, \beta^{p^{nj}} \qquad (j = 0, \cdots, s-1)$$

give rise to the same P of (4.1), we conclude that

$$s\,\psi_2^*(s, p^n) = s\,\psi(s, p^{2n}) = \sum_{s=[e,f]} e\,\psi(e, p^n)\cdot f\,\psi(f, p^n), \tag{4.3}$$

the summation on the right extending over all e, f whose least common multiple equals s.

THEOREM 4.1. *The function* $\Psi(s, p^n) = s\psi(s, p^n)$ *satisfies the identity* (4.3).

It is not difficult to prove this result directly; we need merely (4.2). By the Dedekind inversion formula, an identity $f(s) = g(s)$ is equivalent to

$$\sum_{d \mid s} f(d) = \sum_{d \mid s} g(d).$$

Thus it will suffice to prove—in place of (4.3)—the following:

$$\sum_{s \mid m} \Psi(s, p^{2n}) = \sum_{s \mid m} \sum_{s=[e,f]} \Psi(e, p^n)\Psi(f, p^n). \tag{4.4}$$

Now the summation conditions on the right of (4.4)

$$s \mid m, \qquad [e, f] = s$$

are easily seen to be equivalent to the simpler conditions

$$e \mid m, \qquad f \mid m,$$

e and f independently ranging over the divisors of m. Thus the right side of (4.4) equals

$$\sum_{e \mid m} \Psi(e, p^n) \sum_{f \mid m} \Psi(f, p^n). \tag{4.5}$$

But since $\sum_{e \mid m} \Psi(e, p^n) = p^{nm}$, it follows that (4.4) is an identity, and this in turn implies the truth of (4.3).

As will be shown elsewhere, the identity (4.3) is easily extended—in several directions. For the present we note the generalization

$$\Psi(s, p^{nf}) = \sum_{s=[e_1,\cdots,e_k]} \Psi(e_1, p^{nf_1}) \cdots \Psi(e_k, p^{nf_k}), \tag{4.6}$$

where $f = f_1 + \cdots + f_k$, and the summation extends over all sets $e_1, \cdots, e_k$ with least common multiple equal to s.

5. **The total number of polynomials of degree** s. In the case $k = 1$, the result is immediate: every (primary) polynomial has the form

$$M = x^s + \alpha_1 x^{s-1} + \cdots + \alpha_s, \qquad \alpha_j \text{ in } GF(p^n), \tag{5.1}$$

so that the number is evidently p^{ns}. For $k > 1$, unfortunately an explicit formula (5.1) is not available, and therefore other methods must be used. One simple method is the following.

Define the zeta-function $\zeta^*(w)$ by means of the infinite product

$$\zeta^*(w) = \prod_{P^*}\left(1 + \frac{1}{|P|^w} + \frac{1}{|P^2|^w} + \cdots\right), \tag{5.2}$$

extended over irreducible P^*. Here the "absolute value"

$$|P^e| = p^{nse} \qquad \text{for } s = \deg P;$$

and for arbitrary M, N,

$$|MN| = |M| \cdot |N|,$$

so that generally

$$|M| = p^{nm} \qquad \text{for } m = \deg M. \tag{5.3}$$

Expanding the product on the right side of (5.2), we have

$$\zeta^*(w) = \sum_{M^*} \frac{1}{|M|^w} = \sum_{w=0}^{\infty} \frac{f^*(m)}{p^{nmw}}, \tag{5.4}$$

m

where $f^*(m)$ denotes the number of M^* of degree m. We determine $f^*(m)$ by evaluating $\zeta^*(w)$. Now (5.2) may be put in the form

$$\begin{aligned}\zeta^*(w) &= \prod_{P^*}\left(1 - \frac{1}{|P|^w}\right)^{-1} \\ &= \prod_{s=1}^{\infty}\left(1 - \frac{1}{p^{nsw}}\right)^{-\psi(s,\,p^{nk})}\end{aligned}$$

for by (3.14), $\psi_k^*(s, p^n) = \psi(s, p^{nk})$. Thus

$$\begin{aligned}\log \zeta^*(w) &= \sum_{s=1}^{\infty} \psi(s, p^{nk}) \sum_{t=1}^{\infty} \frac{1}{tp^{nstw}} \\ &= \sum_{m=1}^{\infty} \frac{1}{mp^{nmw}} \sum_{m=st} s\psi(s, p^{nk}) \\ &= \sum_{m=1}^{\infty} \frac{1}{mp^{nmw}}\, p^{nkm} = \sum_{m=1}^{\infty} \frac{1}{mp^{nm(\cancel{k-w})}},\end{aligned}$$

(w−k)

and therefore

$$\zeta^*(w) = \{1 - p^{n(k-w)}\}^{-1}. \tag{5.5}$$

5.4 5.5

Comparing (4.4) with (4.5), we have immediately

THEOREM 5.1. *The number of (factorable) polynomials of degree m containing* x_k^m *is*

$$f^*(m) = p^{nkm}. \tag{5.6}$$

For $k = 1$, *this reduces to the familiar* p^{nm}.

The determination of $f(m)$, the *total* number of M of degree m (not necessarily containing x_k) is somewhat more elaborate. In place of (5.2) we now have

$$\zeta(w) = \prod_P \left(1 + \frac{1}{|P|^w} + \frac{1}{|P^2|^w} + \cdots\right), \tag{5.7}$$

the product now extending over all irreducible P. If we use the fuller notation $\zeta_k^*(w)$ for the $\zeta^*(w)$ of (5.2), it is readily seen that (5.7) implies

$$\zeta(w) = \zeta_k^*(w)\zeta_{k-1}^*(w) \cdots \zeta_1^*(w). \tag{5.8}$$

On the other hand, (5.4) becomes

$$\zeta(w) = \sum_M \frac{1}{|M|^w} = \sum_{m=0}^{\infty} \frac{f(m)}{p^{nmw}}. \tag{5.9}$$

Therefore by (5.5), (5.8), (5.9), we have

$$\sum_{m=0}^{\infty} \frac{f(m)}{p^{nmw}} = \{(1 - p^{n(1-w)})(1 - p^{n(2-w)}) \cdots (1 - p^{n(k-w)})\}^{-1}. \tag{5.10}$$

Now by a familiar expansion,

$$\{(1 - qt)(1 - q^2t) \cdots (1 - q^kt)\}^{-1} = \sum_{j=0}^{\infty} \frac{(q^k - 1) \cdots (q^{k+j-1} - 1)}{(q - 1)\cdots(q^j - 1)} q^j t^j.$$

In this identity put $q = p^n$, $t = p^{-nw}$; thus we have

THEOREM 5.2. *The total number of (primary factorable) polynomials of degree m is determined by*

$$f(m) = \frac{(p^{nk} - 1) \cdots (p^{n(k+m-1)} - 1)}{(p^n - 1) \cdots (p^{nm} - 1)} p^{nm}. \tag{5.11}$$

Here again for $k = 1$, the formula reduces to the familiar p^{nm}.

By use of the functions ζ and ζ^*, it is easy to carry over to the case of k indeterminates a number of theorems[5] on arithmetic functions known in the case $k = 1$. However, we shall not take the space to develop these formulas.

6. **The product of polynomials of fixed degree.**[6] We now seek an expression for

$$F(m) = \prod_{\deg M = m} M, \tag{6.1}$$

the product of *all* (primary factorable) M of degree m. Here again the simple method used in the case $k = 1$ is not applicable.

[5] See L. Carlitz, American Journal of Mathematics, vol. 54 (1932), pp. 39–50; Bulletin of the American Mathematical Society, vol. 37 (1932), pp. 736–744. The latter paper will be referred to as B.

[6] For $k = 1$, see B, p. 742.

If P denotes a typical irreducible of degree s, put

$$M = P^e A, \qquad P \nmid A, \tag{6.2}$$

so that P^e is the highest power of P dividing M. For fixed P and e, what is the number of polynomials A of degree $m - es$ not divisible by P? It is easily seen that this number is

$$\varphi_{m-es}(P) = \begin{cases} f(m - es) \text{ for } m - es < s, \\ f(m - es) - f(m - (e + 1)s) \text{ for } m - es \geqq s, \end{cases} \tag{6.3}$$

where $f(m)$ is defined by (5.11). But by (6.2) it is evident that

$$\prod_{\deg M = m} M = \prod_{P,\, e} P^{e\varphi_{m-es}(P)}, \tag{6.4}$$

the product on the right extending over all irreducible P of degree s, and all positive e such that $es \leqq m$. But the right member of (6.4) equals

$$\prod_P P^{\sum_e e\varphi_{m-es}(P)},$$

the summation in the exponent extending over all $e \leqq m/s$. By (6.3),

$$\begin{aligned} \sum e\varphi_{m-es}(P) &= \{f(m - s) - f(m - 2s)\} \\ &\quad + 2\{f(m - 2s) - f(m - 3s)\} + \cdots + rf(m - rs) \\ &= f(m - s) + f(m - 2s) + \cdots + f(m - rs), \end{aligned} \tag{6.5}$$

where $r = [m/s]$, the greatest integer $\leqq m/s$. Thus by (6.4) and the definition of $\Theta(s)$,

$$\prod_{\deg M = m} M = \prod_{s=1}^{m} \{\Theta(s)\}^E, \tag{6.6}$$

where E denotes the sum (6.5).

Now on the other hand, by (3.7)

$$\begin{aligned} \prod_{j=1}^{m} (D^j)^{f(m-j)} &= \prod_{es \leqq m} \{\Theta(s)\}^{f(m-es)} \\ &= \prod_{s=1}^{m} \{\Theta(s)\}^{\sum_e f(m-es)}. \end{aligned}$$

Comparison with (6.6) and (6.1) shows at once that

$$F(m) = \prod_{j=1}^{m} (D^j)^{f(m-j)}. \tag{6.7}$$

THEOREM 6.1. *The product of all primary factorable polynomials of degree m is given by* (6.7).

As for ΠM^*, it is not difficult to prove, in exactly the same manner that (6.7) was derived, the following

THEOREM 6.2. *The product of those polynomials of degree m that contain* x_k^m *is*

$$F^*(m) = \prod_{j=1}^{m} \left\{ \frac{D^j(x_0 x_1 \cdots x_k)}{D^j(x_0 \cdots x_{k-1})} \right\}^{p^{nk(m-j)}}. \tag{6.8}$$

Comparison of the degree of the two members of (6.8) leads directly to (5.6). Similarly from (6.7) it is possible to derive (5.11), but this is somewhat less immediate. The derivation depends on logarithmic differentiation of the identity (5.10).

7. **The least common multiple of polynomials of fixed degree.**[7] Here the method used in the case $k = 1$ carries over. Thus from

$$D^m = \prod_{s \mid m} \Theta(s),$$

it follows that

$$\begin{aligned} D^1 D^2 \cdots D^m &= \prod_{j=1}^{m} \prod_{s \mid j} \Theta(s), \\ &= \prod_{cs \leq m} \Theta(s) = \prod_{s=1}^{m} \{\Theta(s)\}^{[m/s]}. \end{aligned} \tag{7.1}$$

On the other hand, if $L(m)$ is the least common multiple of *all* polynomials of degree m, it is evident that, for P irreducible of degree s, the highest power of P dividing $L(m)$ is precisely $[m/s]$. Therefore

$$L(m) = \prod_{s=1}^{m} \left\{ \prod_{\deg P = s} P \right\}^{[m/s]} = \prod_{s=1}^{m} \{\Theta(s)\}^{[m/s]}.$$

7.1
Comparison with (6.1) leads immediately to

THEOREM 7.1. *The least common multiple of all factorable polynomials of degree m is*

$$L(m) = D^1 D^2 \cdots D^m. \tag{7.2}$$

Similar reasoning yields

THEOREM 7.2. *The least common multiple of polynomials of degree m that contain x is*

x_k^m

$$L^*(m) = \prod_{j=1}^{m} \frac{D^j(x_0 x_1 \cdots x_k)}{D^j(x_0 \cdots x_{k-1})}.$$

8. **Concluding remark.** According to Theorem 5.1, $f^*(m) = p^{nkm}$; in other words, the number of polynomials $M^*(x_1 \cdots x_k)$ with coefficients in $GF(p^n)$ is identical with the number of $M(x)$ with coefficients in $GF(p^{nk})$. According to the last part of Theorem 3.3, $\psi_k^*(s, p^n) = \psi(s, p^{nk})$, that is, the number of irreducible $P^*(x \cdots x)$ with coefficients in $GF(p^n)$ is equal to the number of $P(x)$ with coefficients in $GF(P^{nk})$. Thus the question arises whether it is possi-

[7] For $k = 1$, see B, p. 742.

ble to set up a correspondence between $M(x)$ in $GF(p^{nk})$ and $M^*(x_1 \cdots x_k)$ in $GF(p^n)$ in such a way that an irreducible $P(x)$ will correspond to an irreducible $P^*(x_1 \cdots x_k)$, and so that all our theorems will follow immediately from the case $k = 1$. As we shall now indicate, this does not seem to be the case.

An irreducible $P^*(x, y)$ according to Theorem 3.1 is of the form

$$\prod_{j=0}^{s-1} (\alpha^{p^{nj}} + \beta^{p^{nj}} x + y), \tag{8.1}$$

where $e = \deg \alpha$, $f = \deg \beta$, $s = [e, f]$. Similarly an irreducible $P(y)$ with coefficients in $GF(p^{2n})$ is of the form

$$\prod_{j=0}^{s-1} (\gamma^{p^{2nj}} + y) \qquad \gamma \text{ in } GF(p^{2ns}), \tag{8.2}$$

where now s equals the degree of γ relative to $GF(p^{2n})$.

We now attempt to identify (8.1) with (8.2).[8] Thus the two sets of quantities

$$\alpha^{p^{nj}} + \beta^{p^{nj}} x, \qquad \gamma^{p^{2nj}} \qquad (j = 1, \cdots, s)$$

must be identical, except for order. But it is easily seen in particular cases that this is impossible. For example, for $p^n = 2 = s$, let

$$\begin{aligned} P^*(x, y) &= (\vartheta + \vartheta^2 x + y)(\vartheta^2 + \vartheta^4 x + y) \\ &= y^2 + (x + 1)y + x^2 + x + 1, \end{aligned}$$

where $\vartheta^2 + \vartheta + 1 = 0$, so that ϑ defines the $GF(2^2)$. Then $P(y) = (\gamma + y)(\gamma^4 + y)$, whence

$$\gamma = \vartheta + \vartheta^2 x, \qquad \gamma^4 = \vartheta^2 + \vartheta x.$$

But these equations imply

$$\vartheta + \vartheta^2 x^4 = \vartheta^2 + \vartheta x.$$

Throwing out $x = 1$, since it leads to $P(y) = (1 + y)^2$, we have

$$(x + 1)^3 = \vartheta + 1. \tag{8.3}$$

On the other hand, x is in $GF(2^4)$, so that

$$(x + 1)^{15} = 1;$$

combining this with (8.3) leads to $\vartheta = 0$.

INSTITUTE FOR ADVANCED STUDY AND DUKE UNIVERSITY.

[8] I am indebted to B. P. Gill for this suggestion.

EQUIVALENCE OF MULTILINEAR FORMS SINGULAR ON ONE INDEX

By Rufus Oldenburger

1. **Introduction.** Any p-way matrix $A = (a_{ij\cdots k})$ of order n can be "factored" in the form

$$A = \left(\sum_{\alpha=1}^{h} a_{\alpha i}\, b_{\alpha j} \cdots d_{\alpha k}\right) \qquad (i, j, \cdots, k = 1, \cdots, n), \tag{1'}$$

where $h \leqq n^{p-1}$. Hitchcock,[1] using the polyadic point of view, has determined minimum values of h for some given numerical values of n and p. The representation (1′) implies that *any multilinear form*

$$F = a_{ij\cdots k}\, x_i y_j \cdots z_k \qquad (i, \cdots, k = 1, \cdots, n)$$

(repeated indices indicate summation) *is equivalent under transformations*

$$x'_\alpha = a_{\alpha i} x_i\,, \tag{2$_1$}$$

$$y'_\beta = b_{\beta j} y_j\,, \tag{2$_2$}$$

$$\cdots \qquad \cdots$$

$$z'_\gamma = d_{\gamma k} z_k\,, \tag{2$_p$}$$

to the form

$$R = x'_\alpha y'_\alpha \cdots z'_\alpha \qquad (\alpha = 1, \cdots, h),$$

where $h \leqq n^{p-1}$ and the transformations $(2_1), \cdots, (2_p)$ are not necessarily non-singular.

We shall say that the matrix $(a_{ij\cdots k}\, a_{\alpha i})$ of the form F' obtained from F by applying the transformation $x_i = a_{\alpha i} x'_\alpha$ to F, where $(a_{\alpha i})$ is non-singular, *is equivalent to* $(a_{ij\cdots k})$; we shall also say that F' is equivalent to F. If the 2-way matrices $(a_{\alpha i}), \cdots, (d_{\alpha k})$ of (1′) are all singular on their columns (α being taken as the row index in these matrices), the matrix A is equivalent to a matrix of lower order of the form (1′), where at least one of the matrices $(a_{\alpha i}), \cdots, (d_{\alpha k})$ is non-singular on its columns. The number h of (1′) is then between the limits $n \leqq h \leqq n^{p-1}$. In another paper[2] the author treated the special case where h takes the minimum value n. He obtained necessary and sufficient conditions for the factorability of a matrix A into the form (1′), where the matrices $(a_{\alpha i}), \cdots, (d_{\alpha k})$ are all non-singular. The method of

Received January 24, 1936; in revised form, June 11, 1936.

[1] F. L. Hitchcock, *A new method in the theory of quantics*, Journal of Mathematics and Physics, vol. 8 (1929), p. 83.

[2] R. Oldenburger, *Non-singular multilinear forms and certain p-way matrix factorizations*, Transactions of the American Mathematical Society, vol. 39 (1936), pp. 422–455. This paper will be denoted by N. S.

treatment extended at once to the case where all but one of the matrices $(a_{\alpha i})$, $\cdots$, $(d_{\alpha k})$ are non-singular. In the case where these matrices are all non-singular the matrices of type A are at once equivalent to each other for given n, p, and their associated multilinear forms are equivalent to the unique canonical form R.

In the present paper, the treatment of equivalence of matrices of type A, where $h = n$, and one of the matrices $(a_{\alpha i})$, $\cdots$, $(d_{\alpha k})$ is singular, is completed. *Necessary and sufficient conditions are obtained for the equivalence of two such matrices for given n, p, and some of the associated canonical forms are derived.* In contrast to the 2-way case the number of canonical forms for given n, p, where $p \geqq 3$, $n \geqq 4$, is *infinite*.

In the author's Transactions paper the matrices of type (1′), where $h = n$, and $(a_{\alpha i})$, $\cdots$, $(d_{\alpha k})$ are non-singular, were said to be non-singular. The associated forms were also said to be non-singular. The matrix (1′) and its associated multilinear form is said to be *singular* and of rank[3] r on the index i, and non-singular on j, $\cdots$, k if $(a_{\alpha i})$ is of rank r, and $(b_{\alpha j})$, $\cdots$, $(d_{\alpha k})$ are non-singular.

The development of the present paper is based on the property that two given singular forms, having the same p, n, and r, are equivalent if and only if invariant factors associated with reduced forms are equivalent under certain non-singular linear transformations. These invariant factors are generalizations of the invariant factors[4] of $(\rho A + \sigma B)$, where A, B are two-way matrices, and ρ, σ are parameters. It will be no restriction on the generality of the method to take $p = 3$. The associated forms are then trilinear.

The theory developed here holds for any field of numbers.

2. Invariant factors. Let a given 3-way matrix singular on one index i be denoted by

$$D = \left(\sum_{\alpha=1}^{n} a_{\alpha i} b_{\alpha j} c_{\alpha k}\right); \qquad (i = 1, \cdots, r < n; j, k = 1, \cdots, n).$$

The range chosen for i does not restrict the generality, since $(a_{\alpha i})$, $\alpha, i = 1, \cdots, n$, of rank $r < n$ can be reduced by multiplication on the right with a non-singular matrix to a matrix $(a_{\alpha i})$ as given in D bordered by zeros. The matrix $(a_{\alpha i})$ is equivalent under multiplication on the right and rearrangement of the rows to

$$G = \begin{pmatrix} I \\ A' \end{pmatrix},$$

where I is a Kronecker delta of order r. Let rearrangement of the rows in $(a_{\alpha i})$ be simultaneously accompanied by a similar rearrangement of the rows

[3] For a treatment of ranks of this type see the author's paper *Composition and rank of n-way matrices and multilinear forms*, Annals of Mathematics, vol. 35 (1934), pp. 622–657.

[4] For the definition of these invariant factors see L. E. Dickson, *Modern Algebraic Theories*, 1926, p. 104.

of $(b_{\alpha j})$, $(c_{\alpha k})$. The matrices $(b_{\alpha j})$, $(c_{\alpha k})$ are equivalent under multiplication on the right with reciprocals to $(\delta_{\alpha\alpha'})$, $(\delta_{\alpha\alpha''})$, respectively, where $(\delta_{\alpha\alpha'})$, $(\delta_{\alpha\alpha''})$ are Kronecker deltas of order n. The matrix D is hence equivalent to

$$E = \left(\sum_{\alpha=1}^{n} a_{\alpha i}\delta_{\alpha\alpha'}\delta_{\alpha\alpha''}\right),$$

where $(a_{\alpha i})$ is of the form G. Let $A' = (a'_{ji})$. The form associated with E is

$$F = \sum_{\alpha=1}^{r} x_\alpha y_\alpha z_\alpha + \sum_{j=r+1}^{n} \sum_{i=1}^{r} a'_{ji} x_i y_j z_j. \tag{1}$$

Elsewhere the author has discussed[5] the *i-invariant factors* of a matrix (a_{ijk}) of order n and its associated form, which are quotients of the type

$$I_{n-j+1} = \frac{G_j}{G_{j-1}},$$

where, except for a constant factor, G_j is the g.c.d. of the j-th order determinant minors of the i-characteristic matrix

$$M(i) = (\rho_1 a_{1jk} + \cdots + \rho_n a_{njk}).$$

The ρ's in $M(i)$ are parameters. To obtain the G_j, it is necessary to factor the j-th order determinant minors of this matrix in the given field. It is to be noted that since these determinants are homogeneous polynomials in $\rho_1, \cdots, \rho_n$, if $n > 2$, they cannot be broken up in general into linear factors in any field or even when the coefficients belong to a linear algebra.[6] The author proved that transformations

$$x_i = b_{im} x'_m$$

on $a_{ijk} x_i y_j z_k$ correspond to the transformations

$$\rho_i = b_{im} \rho'_m$$

on the ρ's. The author also proved that under non-singular linear transformations on y_j and z_k the G_j (hence G_j/G_{j-1}) are relative invariants. The i-characteristic matrix $M(i)$ of F is the diagonal matrix

$$M(i) = \begin{bmatrix} \rho_1 & & & & & & \\ & \rho_2 & & & & 0 & \\ & & \ddots & & & & \\ & & & \rho_r & & & \\ & 0 & & & (\rho_1 a'_{r+1,1} + \cdots + \rho_r a'_{r+1,r}) & & \\ & & & & & \ddots & \\ & & & & & & (\rho_1 a'_{n1} + \cdots + \rho_r a'_{nr}) \end{bmatrix}.$$

[5] N. S., §6.

[6] Footnote p. 432, N. S., is incorrect.

Since $M(i)$ is diagonal, we have

LEMMA 1. *The i-invariant factors of the form F given in* (1) *different from constants are products of the linear forms*

$$\rho_1, \cdots, \rho_r, L_{r+1} = \sum_{i=1}^{r} a'_{r+1,i}\rho_i, \cdots, L_n = \sum_{i=1}^{r} a'_{ni}\rho_i, \tag{2}$$

each of these expressions occurring exactly once in the invariant factors.

Since the j- and k-ranks of F are equal to n, none of these linear forms vanishes identically. The matrix of these linear forms[7] is exactly G.

In the following we shall use $I_1, \cdots, I_s$ to denote i-invariant factors of F which are not constants.

THEOREM 1. *Two forms F and F' of type* (1) *are equivalent if and only if they have the same number of invariant factors distinct from constants, and the i-invariant factors $I_1, \cdots, I_s$ of F are simultaneously equivalent to $k_1I'_1, \cdots, k_sI'_s$, where $k_1, \cdots, k_s$ are constants, and $I'_1, \cdots, I'_s$ are the corresponding i-invariant factors of F'.*

The necessity of the conditions of the theorem follow from the above remarks on the manner in which i-invariant factors of a form F are changed when transformations are made on F.

To prove the sufficiency of the conditions, let

$$I_\xi = M_1^\xi(\rho_1, \cdots, \rho_r) \cdots M_{i_\xi}^\xi(\rho_1, \cdots, \rho_r),$$

where $M_1^\xi(\rho_1, \cdots, \rho_r), \cdots, M_{i_\xi}^\xi(\rho_1, \cdots, \rho_r)$ are linear forms in $\rho_1, \cdots, \rho_r$; similarly let

$$I'_\xi = M_1^{\xi\prime}(\rho'_1, \cdots, \rho'_r) \cdots M_{i_\xi}^{\xi\prime}(\rho'_1, \cdots, \rho'_r).$$

Assume that there exist non-singular linear transformations

$$\rho_m = b_{mi}\,\rho'_i \qquad (i, m = 1, \cdots, r), \tag{3}$$

and k_ξ, $\xi = 1, \cdots, s$, such that

$$I'_\xi = k_\xi M_1^\xi(b_{1i}\,\rho'_i, \cdots, b_{ri}\,\rho'_i) \cdots M_{i_\xi}^\xi(b_{1i}\,\rho'_i, \cdots, b_{ri}\,\rho'_i). \tag{4}$$

Now (4) implies that there exist constants $C_{\xi 1}, \cdots, C_{\xi i_\xi}$, where $C_{\xi 1} \cdots C_{\xi i_\xi} = k_\xi$, such that the linear forms

$$C_{\xi 1}\, M_1^\xi(b_{1i}\,\rho'_i, \cdots, b_{ri}\,\rho'_i), \cdots, C_{\xi i_\xi}\, M_{i_\xi}^\xi(b_{1i}\,\rho'_i, \cdots, b_{ri}\,\rho'_i)$$

are identically equal in some order to

$$M_1^{\xi\prime}(\rho'_1, \cdots, \rho'_r), \cdots, M_{i_\xi}^{\xi\prime}(\rho'_1, \cdots, \rho'_r).$$

By the lemma,

$$M_1^\xi(\rho_1, \cdots, \rho_r), \cdots, M_{i_\xi}^\xi(\rho_1, \cdots, \rho_r) \qquad (\xi = 1, \cdots, s)$$

[7] That is, the matrix of the bilinear form $\sigma_1\rho_1 + \cdots + \sigma_r\rho_r + \sigma_{r+1}\sum_{i=1}^{r} a'_{r+1,i}\rho_i + \cdots + \sigma_n \sum_{i=1}^{r} a'_{ni}\rho_i$ is G.

are equal in some order to the diagonal elements of the i-characteristic matrix $M(i)$ of F given above. Let F' be given by

$$F' = \sum_{\alpha=1}^{r} x'_\alpha y'_\alpha z'_\alpha + \sum_{j=r+1}^{n} \sum_{i=1}^{r} a''_{ji} x'_i y'_j z'_j .$$

Now $M_1^{\xi\prime}(\rho'_1, \cdots, \rho'_r), \cdots, M_{i_\xi}^{\xi\prime}(\rho'_1, \cdots, \rho'_r)$, $\xi = 1, \cdots, s$, are equal in some order to the diagonal elements of the i-characteristic matrix of F, which are given by

$$\rho'_1, \cdots, \rho'_r, L'_{r+1} = \sum_{i=1}^{r} a''_{r+1,i} \rho'_i, \cdots, L'_n = \sum_{i=1}^{r} a''_{ni} \rho'_i .$$

It follows by (2) and (3) that there exist constants C_i, $i = 1, \cdots, n$, such that

$$\sum_{n=1}^{r} C_1 b_{1m} \rho'_m, \cdots, \sum_{m=1}^{r} C_r b_{rm} \rho'_m,$$

$$C_{r+1} \sum_{m,i=1}^{r} a'_{r+1,m} b_{mi} \rho'_i, \cdots, C_n \sum_{m,i=1}^{r} a'_{n.m} b_{mi} \rho'_i$$

are equal in some order to

$$\rho'_1, \cdots, \rho'_r, L'_{r+1}, \cdots, L'_n .$$

This implies that there exist non-singular transformations (3) and

(5) $$C_i \sigma'_i = \sigma_j \qquad (i, j = 1, \cdots, n; i \text{ not summed}),$$

where there is one j for every i, such that the bilinear form

(6) $$\sigma_1 \rho_1 + \cdots + \sigma_r \rho_r + \sigma_{r+1} L_{r+1} + \cdots + \sigma_n L_n$$

transforms into the form

(7) $$\sigma'_1 \rho'_1 + \cdots + \sigma'_r \rho'_r + \sigma'_{r+1} L'_{r+1} + \cdots + \sigma'_n L'_n .$$

Now (6) and (7) become F and F' if we make the substitutions

$$\rho_i = x_i, \qquad \rho'_i = x'_i, \qquad \sigma_i = y_i z_i, \qquad \sigma'_i = y'_i z'_i \quad (i \text{ not summed}).$$

By (3) and (5) F becomes F' if we let

(8) $$y_i z_i = C_j y'_j z'_j \qquad (i, j = 1, \cdots, n, \text{ not summed}),$$

(9) $$x_m = b_{mi} x'_i \qquad (m, i = 1, \cdots, r).$$

Equation (8) is equivalent to

(10) $$y_i = C_j y'_j, \qquad z_i = z'_j \quad (i, j = 1, \cdots, n, \text{ not summed}).$$

Hence F is equivalent to F' under the non-singular linear transformations (9) and (10).

Since the i-invariant factor I_1 of F contains the factors $\rho_1, \cdots, \rho_r$, the

matrix of the linear factors of I_1, which will be called the matrix of I_1, can be written in the form

$$\begin{pmatrix} I \\ K_1 \end{pmatrix},$$

where I is a Kronecker delta of order r. Let matrices of the i-invariant factors $I_1, \cdots, I_s$ of F be denoted by

$$\begin{pmatrix} I \\ K_1 \end{pmatrix}, \quad K_2, \quad \cdots, \quad K_s \tag{11}$$

respectively, and matrices of the corresponding i-invariant factors $I'_1, \cdots, I'_s$ of form F' by

$$\begin{pmatrix} I \\ K'_1 \end{pmatrix}, \quad K'_2, \quad \cdots, \quad K'_s \tag{12}$$

respectively. Let a non-singular matrix J which permutes the rows or columns of a 2-way matrix A under the operation JA or AJ be called a *permutation matrix.*

Equations (3) and (5) imply the

COROLLARY. *Two forms F and F' of type* (1), *which have the same number of i-invariant factors distinct from constants and corresponding i-invariant factors are of the same degree, are equivalent if and only if there exist non-singular permutation and diagonal matrices J_i, α_i, and a non-singular matrix X such that*

$$\alpha_1 J_1 \begin{pmatrix} I \\ K_1 \end{pmatrix} X = \begin{pmatrix} I \\ K'_1 \end{pmatrix},$$

$$\alpha_i J_i K_i X = K'_i \qquad (i = 2, \cdots, s),$$

where $I, K_1, \cdots, K_s, K'_1, \cdots, K'_s$ are as given in (11) *and* (12).

We have now

LEMMA 2. *If there exists a diagonal matrix*

$$\alpha = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix},$$

where α_1 is a minor of order r, and a non-singular 2-way matrix X such that $\alpha_1 IX = I$, where I is a Kronecker delta, then X is a diagonal matrix and

$$\alpha \begin{pmatrix} I \\ A \end{pmatrix} X = \begin{pmatrix} I \\ \alpha_2 AX \end{pmatrix}.$$

If

$$\alpha_1 = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_r \end{pmatrix}$$

and $\alpha_1 IX = I$, then

$$X = \begin{pmatrix} \frac{1}{a_1} \cdot & & 0 \\ & \cdot & \\ & & \cdot \frac{1}{a_r} \\ 0 & & \end{pmatrix}.$$

Evidently,

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} X = \begin{pmatrix} \alpha_1 \, IX \\ \alpha_2 \, AX \end{pmatrix}.$$

Let α and β denote non-singular diagonal 2-way matrices, and let X be a non-singular 2-way matrix. If elements in two matrices are in corresponding position, these elements will be said to *correspond*. We similarly define corresponding rows and columns. We now have the following results.

LEMMA 3. *If $\alpha A \beta = A'$, the vanishing elements of A and A' correspond.*

THEOREM 2. *If $AX = A'$, the minors composed of corresponding rows of A and A' are of the same rank.*

Theorem 2 implies the

COROLLARY. *The matrix*

$$\begin{pmatrix} I \\ K_1 \\ K_p \\ K_s \\ \vdots \\ K_t \end{pmatrix}$$

of the i-invariant factors $I_1, I_p, I_s, \cdots, I_t$ of F has the same number of non-singular minors of maximum order r as the matrix

$$\begin{pmatrix} \alpha_1 J_1 & & & & \\ & \alpha_p J_p & & 0 & \\ & & \alpha_s J_s & & \\ & & & \ddots & \\ 0 & & & & \alpha_t J_t \end{pmatrix} \begin{pmatrix} I \\ K_1 \\ K_p \\ K_s \\ \vdots \\ K_t \end{pmatrix},$$

where $\alpha_1, \alpha_p, \cdots, \alpha_t$ and $J_1, J_p, \cdots, J_t$ are diagonal and permutation matrices of the same orders as there are rows in $\begin{pmatrix} I \\ K_1 \end{pmatrix}$, $K_p, \cdots, K_t$ respectively.

3. **The case** $r = n - 1$. Using the results of §2 we shall prove

THEOREM 3. *If $r = n - 1$, the form F is equivalent to a canonical form of the type*

$$C = \sum_{\alpha=1}^{n-1} x_\alpha y_\alpha z_\alpha + y_n z_n \left(\sum_{\alpha=1}^{\sigma} x_\alpha \right),$$

where σ is unique and $\leqq n - 1$.

The form F given in (1) is now

$$F = \sum_{\alpha=1}^{n-1} x_\alpha y_\alpha z_\alpha + \sum_{i=1}^{n-1} a'_{ni} x_i y_n z_n. \tag{13}$$

If the σ elements $a'_{n,p}, a'_{n,q}, \cdots, a'_{n,s} \neq 0$, and all other elements in the array $a'_{n,1}, \cdots, a'_{n,n-1}$ are zero, make the substitutions on the x's, y's and z's in (13) to give a new form of type (13), where

$$a'_{n,1}, \cdots, a'_{n,\sigma} \neq 0; a'_{n,\sigma+1}, \cdots, a'_{n,n-1} = 0. \tag{14}$$

Assume that F is of type (13), where (14) is satisfied. Making the transformations

$$x_1 = \frac{x'_1}{a'_{n1}}, \cdots, x_\sigma = \frac{x'_\sigma}{a'_{n\sigma}}, y_1 = a'_{n1} y'_1, \cdots, y_\sigma = a'_{n\sigma} y'_\sigma \tag{15}$$

on F, we obtain C. The transformations used to obtain C are non-singular.

To prove that σ is unique, we consider the case where C has only one i-invariant factor, I_1, distinct from a constant. By Lemma 1

$$I_1 = \rho_1 \cdots \rho_{n-1} \left(\sum_{\alpha=1}^{\sigma} \rho_\alpha \right),$$

whose matrix is

$$\begin{pmatrix} I \\ K \end{pmatrix} = \begin{pmatrix} 1 & \cdot & \cdot & 0 \\ 0 & & \cdot & 1 \\ 1 \cdots & 1 & 0 \cdots & 0 \end{pmatrix},$$

where there are $\sigma \geqq 2$ unit elements in the last row, and there are $n - 1$ columns. There are $\sigma + 1$ non-singular $(n - 1)$ order minors in the above matrix, whence σ is invariant by the corollaries to Theorems 1 and 2.

If C has more than one i-invariant factor distinct from a constant, by the lemma these invariant factors are

$$I_1 = \rho_1 \cdots \rho_{n-1}, \qquad I_2 = \sum_{\alpha=1}^{\sigma} \rho_\alpha,$$

whence, since I_2 divides I_1, we have $\sigma = 1$. In this case there is only one non-singular minor of order $(n - 1)$ of the matrix of I_1, which property is invariant by the corollaries of Theorems 1 and 2.

4. **The equivalence of forms with** $r = n - 2$. For forms with 2 or 3 i-invariant factors we have

THEOREM 4. *A trilinear form F with $r = n - 2$, and*

(a) *two i-invariant factors distinct from constants, one of which is of degree 2, and the other of degree $\geqq 2$, is equivalent to*

$$\sum_{\alpha=1}^{n-2} x_\alpha y_\alpha z_\alpha + x_1 y_{n-1} z_{n-1} + x_2 y_n z_n;$$

(b) *two i-invariant factors distinct from constants, one of which is of degree* 1 *and the other of degree* $\geqq 1$, *is equivalent to*

$$\sum_{\alpha=1}^{n-2} x_\alpha y_\alpha z_\alpha + \left(\sum_{\alpha=1}^{\rho} x_\alpha\right)(y_{n-1}z_{n-1} + y_n z_n) \qquad (2 \leqq \rho \leqq n-2),$$

or

$$\sum_{\alpha=1}^{n-2} x_\alpha y_\alpha z_\alpha + y_{n-1}z_{n-1}\left(\sum_{\alpha=1}^{\sigma} x_\alpha\right) + y_n z_n x_{\sigma+1} \qquad (2 \leqq \sigma \leqq n-3),$$

(c) *three i-invariant factors distinct from constants are equivalent to*

$$\sum_{\alpha=1}^{n-2} x_\alpha y_\alpha z_\alpha + x_1(y_{n-1}z_{n-1} + y_n z_n).$$

These canonical forms are not equivalent.

For the sake of brevity we omit the proof which involves the use of lemmas and theorems proved in the preceding section.

For the case not treated in Theorem 4 the form F has one i-invariant factor distinct from a constant. It can be proved that two such forms are equivalent if and only if certain rational functions of the coefficients of one form are equal to corresponding rational functions of the coefficients of the other. For a given $n \geqq 4$, the canonical forms are unlimited in number so that there are an unlimited number of sets of equivalent forms.

5. **Necessary and sufficient condition for equivalence.** To distinguish between the canonical forms of part (b) of Theorem 4 we note that the matrix of the i-invariant factors of the first form is

$$\begin{pmatrix} I \\ K_1 \\ K_2 \end{pmatrix} = \begin{pmatrix} & I & \\ 1 \cdots & 10 & \cdots 0 \\ 1 \cdots & 10 & \cdots 0 \end{pmatrix},$$

where I is a Kronecker delta of order $n-2$, and there are ρ unit elements in each of the last two rows. The corresponding matrix for the second form is

$$\begin{pmatrix} & I & \\ 1 \cdots & 100 & \cdots 0 \\ 0 \cdots & 010 & \cdots 0 \end{pmatrix},$$

where there are σ unit elements in the next to the last row. Assume that the necessary equivalence condition $\rho = \sigma$ is satisfied. In the first matrix there are $2\rho + 1$ non-singular minors of order $n-2$. In the second there are $\rho + 2$ such minors. Since $\rho > 1$, we have $2\rho + 1 \neq \rho + 2$. By the corollaries to Theorems 1 and 2 the associated forms are not equivalent. We have proved

THEOREM 5. *Let F and F' be two trilinear forms of type* (1) *with $r = n-1$, or with $r = n-2$ but* 2 *or* 3 *i-invariant factors. The forms F and F' are equivalent if and only if*

(a) *they have the same number of i-invariant factors distinct from constants,*

(b) *corresponding i-invariant factors are of the same degree,*

(c) *the matrix of the i-invariant factor I_1 of F and the matrix of the i-invariant factors I_1, I_2 of F have the same numbers of non-singular minors of order r as the corresponding matrices for F'.*

6. **Conclusion.** We have considered all possible cases of trilinear forms where $r = n - 1$ or $n - 2$. The case $r = n - 2$ is typical of the cases where $r < n - 2$.

Let F denote a singular multilinear form

$$\sum_{i_1, \cdots, i_p = 1}^{n} a_{i_1 \cdots i_p} x_{i_1}^{(1)} \cdots x_{i_p}^{(p)}$$

singular on one index i_1 and of rank $r = n - 1$ or $n - 2$ on this index. To obtain the canonical forms to which F is equivalent from the canonical forms of this paper simply replace x_α by $x_\alpha^{(1)}$ and $y_\alpha z_\alpha$ by $x_\alpha^{(2)} \cdots x_\alpha^{(p)}$ for every α. The i-invariant factors for a general F are defined in terms of the space determinant minors of a matrix associated with F.

ARMOUR INSTITUTE OF TECHNOLOGY.

THE QUADRATIC SUBFIELDS OF A GENERALIZED QUATERNION ALGEBRA

By Claiborne G. Latimer

1. **Introduction.** Let $\mathfrak{A}$ be a rational generalized quaternion algebra with the fundamental number d, as defined by Brandt.[1] Every element of $\mathfrak{A}$, not rational, is a root of a quadratic equation with rational coefficients, and hence defines a quadratic field. The question arises as to what quadratic fields are contained in $\mathfrak{A}$. The purpose of this note is to prove the following

Theorem. *Let $\mathfrak{A}$ be a rational generalized quaternion algebra, with the fundamental number d, and let F be a quadratic field. $\mathfrak{A}$ contains a field equivalent to F if and only if*

(a) *F is imaginary when $d > 0$;*

(b) *no rational prime factor of d is the product of two distinct prime ideals in F.*

Hasse proved a theorem on the splitting fields of an algebra which, when properly specialized, is equivalent to the above theorem, his results being in terms of the p-adic extensions of $\mathfrak{A}$ and of F.[2] Our proof is independent of Hasse's and is short and elementary.

2. **Proof of necessary conditions.** Suppose $\mathfrak{A}$ contains F. Let F be defined by $(-\alpha)^{\frac{1}{2}}$, α being an integer with no square factor > 1. If $d > 0$, by the definition of d, $\mathfrak{A}$ contains no element with a negative norm. Hence F is imaginary.

$\mathfrak{A}$ contains an element i such that $i^2 = -\alpha$. Then the trace, or double the scalar part, of i is zero. It may be shown that $\mathfrak{A}$ also contains a non-singular element j, such that the trace of j and the trace of ij are zero. Then $1, i, j, ij$ are linearly independent, and hence form a basis of $\mathfrak{A}$, $j^2 = -\beta \neq 0$, where β is rational, and $ji = -ij$. We shall assume, without loss of generality, that β is a rational integer with no square factor >1.

Let $\alpha = \alpha_1\delta$, $\beta = \beta_1\delta$, where δ is the positive g.c.d. of α and β. Then $d = \pm AB\Delta$ or $d = \pm 2AB\Delta$, where A, B, Δ are certain positive odd divisors of α, β, δ respectively.[3] By the same reference, d is even if and only if

$$(1) \qquad (\alpha_1 + \beta_1)(\beta_1 + \delta)(\delta + \alpha_1)(\alpha_1 + \beta_1 + \delta) \equiv 8 \pmod{16}.$$

Received June 8, 1936.

[1] Brandt, *Idealtheorie in Quaternionenalgebren*, Mathematische Annalen, vol. 99 (1928), p. 9.

[2] Hasse, *Die Struktur der R. Brauerschen algebrenklassengruppe über einem algebraischen Zahlkörper*, Mathematische Annalen, vol. 107 (1933), pp. 731–760; Deuring, *Algebren*, p. 118.

[3] *On the fundamental number of a rational generalized quaternion algebra*, this Journal, vol. 1 (1935), pp. 433–435. This paper will be referred to hereafter as FN.

Let p be a rational prime divisor of d and consider the principal ideal $\{p\}$ in F. If p divides $A\Delta$, it divides the discriminant, $-\alpha$ or -4α, of F and $\{p\}$ is the square of a prime ideal in F. If p divides B, by the definition of B in FN, $-\alpha$ is a quadratic non-residue of p. Hence $\{p\}$ is a prime ideal in F. Suppose $p = 2$. If $\alpha \equiv 1$ or $2 \pmod 4$, 2 is a divisor of the discriminant of F and hence $\{2\}$ is the square of a prime ideal in F. Suppose $\alpha \equiv 3 \pmod 4$, and hence $\alpha_1 + \delta \equiv 0 \pmod 4$. By (1), β is even and $\alpha_1 + \delta \equiv 4 \pmod 8$. Hence $\alpha \equiv 3 \pmod 8$. Since the discriminant of F is $-\alpha$, it follows that $\{2\}$ is a prime in F. This proves that (b) is a necessary condition.

3. **Proof of sufficient conditions.** Suppose the conditions (a) and (b) are satisfied. Let F be defined by $(-\alpha)^{1/2}$, as before.

We shall show that there is an integer β, such that if $\mathfrak{A}_1$ is the algebra with the basis $1, i, j, ij$, where $i^2 = -\alpha$, $j^2 = -\beta$, $ij = -ji$, then the fundamental number of $\mathfrak{A}_1$ is d. It follows that $\mathfrak{A}_1$ is equivalent[4] to $\mathfrak{A}$. Since $\mathfrak{A}_1$ obviously contains a subfield equivalent to F, the same is true of $\mathfrak{A}$.

By the theorem of FN, d contains no square factor >1. Then

$$\alpha = 2^e\mu\alpha'\rho, \qquad d = 2^f\nu D\rho,$$

where $e = 0$ or 1, $f = 0$ or 1, $\mu = \pm 1$, $\nu = \pm 1$, and α', D, ρ are positive odd integers, relatively prime in pairs, which we shall assume for the present are all >1. Let

$$\alpha' = p_1p_2 \cdots p_k, \qquad D = q_1q_2 \cdots q_s, \qquad \rho = \rho_1\rho_2 \cdots \rho_t,$$

where the p_i, q_i, ρ_i are primes. By (b) of the theorem, employing Legendre symbols, we have

$$\left(\frac{-\alpha}{q_i}\right) = -1 \qquad (i = 1, 2, \cdots, s). \tag{2}$$

Let $\beta_1 \equiv d/\rho = 2^f\nu D$. By Dirichlet's theorem on the primes in an arithmetic progression, there is an odd prime P, prime to αd, such that

$$\left(\frac{P}{p_i}\right) = \left(\frac{-\beta_1}{p_i}\right), \quad \left(\frac{P}{\rho_j}\right) = -\left(\frac{-\beta_1}{\rho_j}\right) \quad (i = 1, 2, \cdots, k;\, j = 1, 2, \cdots, t), \tag{3}$$

and such that the residue of P, mod 8, is an arbitrarily chosen odd integer. This residue will be specified in one case later on.

By (3), employing Jacobi symbols, we have

$$\left(\frac{P}{\alpha'}\right) = \left(\frac{-\beta_1}{\alpha'}\right), \quad \left(\frac{P}{\rho}\right) = \left(\frac{-\beta_1}{\rho}\right)(-1)^t. \tag{4}$$

Setting $\alpha_1 \equiv \alpha'\rho$, employing the quadratic reciprocity theorem, and noting that $(\mu/P) = (-1)^h$, $h = \frac{1}{4}(\mu - 1)(P - 1)$, we have

[4] Brandt, loc. cit., p. 12.

$$\left(\frac{-\alpha}{P}\right) = \left(\frac{2^e}{P}\right)\left(\frac{\mu}{P}\right)\left(\frac{-\alpha_1}{P}\right) = \left(\frac{P}{\alpha_1}\right)(-1)^K,$$

$$K \equiv \frac{\alpha_1+\mu}{2}\cdot\frac{P-1}{2} + \frac{e}{8}(P^2-1).$$

From (4) and the last equation, we have

$$(5) \qquad \left(\frac{-\alpha}{P}\right) = \left(\frac{-\beta_1}{\alpha_1}\right)(-1)^{K+t}.$$

Suppose $e = 1$ or $\alpha_1 + \mu \equiv 2 \pmod 4$. Then the residue of P, mod 8, may be chosen so that K is even or odd, the choice being made so that the right member of (5) is unity. Then $-\alpha$ is a quadratic residue of P.

Suppose $e = 0$ and $\alpha_1 + \mu \equiv 0 \pmod 4$. Then $\alpha = \alpha_1\mu \equiv 3 \pmod 4$, K is even and by (5)

$$(6) \qquad \left(\frac{-\alpha}{P}\right) = \left(\frac{\alpha_1}{D}\right)(-1)^{t+L}, \qquad L \equiv \frac{\alpha_1-1}{2}\cdot\frac{\nu+D}{2} + \frac{f}{8}(\alpha_1^2-1).$$

By (2),

$$(-1)^s = \left(\frac{-\mu}{D}\right)\left(\frac{\alpha_1}{D}\right), \qquad \left(\frac{\alpha_1}{D}\right) = (-1)^{s+M}, \qquad M \equiv \frac{D-1}{2}\cdot\frac{\mu+1}{2}.$$

Then by (6),

$$\left(\frac{-\alpha}{P}\right) = (-1)^T, \qquad T \equiv s + t + L + M.$$

As noted in the last paragraph of FN, d is positive or negative according as it is the product of an odd or an even number of primes. Hence $s + t \equiv \frac{1}{2}(\nu + 1) \pmod 2$ or $s + t \equiv \frac{1}{2}(\nu - 1) \pmod 2$ according as $f = 0$ or $f = 1$. Suppose $f = 1$. Then by (b) of the theorem, $\alpha \equiv 3 \pmod 8$ and $1 \equiv \frac{1}{8}(\alpha_1^2 - 1) \pmod 2$. Hence for $f = 0$ or $f = 1$, $s + t + \frac{1}{8}f(\alpha_1^2 - 1) \equiv \frac{1}{2}(\nu + 1) \pmod 2$. Since $\alpha_1 \equiv -\mu \pmod 4$, it follows that

$$T \equiv \frac{\nu+1}{2} - \frac{\mu+1}{2}\cdot\frac{\nu+D}{2} + M \equiv \frac{\nu+1}{2}\cdot\frac{1-\mu}{2} \pmod 2.$$

If $\nu = 1$, by (a) of the theorem, $\mu = 1$. Hence in every case T is even. Then $-\alpha$ is a quadratic residue of P.

Let $\beta \equiv \beta_1 P = 2^f \nu DP$. Then, employing (2) and (3), we have

(i) $-\beta$ is a quadratic residue of every prime factor of α',
(ii) $-\beta$ is a quadratic non-residue of every prime factor of ρ,
(iii) $-\alpha$ is a quadratic non-residue of every prime factor of D,
(iv) $-\alpha$ is a quadratic residue of P.

We have assumed heretofore that α', ρ, D are greater than 1. If $\alpha' = 1$ or $\rho = 1$, our former definition of P is without meaning. If $\alpha' = \rho = 1$, i.e.,

$\alpha = \pm 2$ or 1, let P be any prime in the form $8n + 1$. If $\alpha' = 1$, $\rho > 1$, let P be defined as above, except that the symbols in (3) involving the p_i are ignored; similarly for the case $\alpha' > 1$, $\rho = 1$. If $D = 1$, let P be defined as above. Then P is defined in every case and the conditions (i) to (iv) are satisfied, some perhaps vacuously.

Let $\mathfrak{A}_1$ be the algebra with the basis $1, i, j, ij$, where $i^2 = -\alpha, j^2 = -\beta, ij = -ji$. α and β have no common odd prime factor. Hence, by the theorem of FN, the fundamental number of $\mathfrak{A}_1$ is $d_1 = \pm AB$ or $d_1 = \pm 2AB$, where A, B are the least positive odd divisors of α, β respectively such that $-\beta$, $-\alpha$ are quadratic residues of α/A, β/B respectively. By the conditions (i) to (iv), $A = \rho$, $B = D$. By the theorem of FN, $d_1 > 0$ if and only if $\alpha > 0$, $\beta > 0$. β has the same sign as d and by (a) of the theorem, if $d > 0$, then $\alpha > 0$. Hence d_1 and d have the same sign and they are equal or one of them is the double of the other. But we have seen that the sign of a fundamental number is determined by the parity of the number of primes dividing it. Hence $d_1 = d$. By the fourth and fifth sentences of this paragraph, it follows that $\mathfrak{A}$ contains a field equivalent to F and the theorem is proved.

UNIVERSITY OF KENTUCKY.

SEMI-CLOSED SETS AND COLLECTIONS

By G. T. Whyburn

1. A set K in a metric space S will be said to be *semi-closed* provided each component of K is closed and any convergent sequence of components of K whose limit set intersects $S - K$ converges to a single point of $S - K$.

Similarly, a collection G of disjoint sets is said to be semi-closed if each set of G is closed and any convergent sequence of sets of G whose limit set intersects $S - G^*$ converges to a single point of $S - G^*$, where G^* denotes the point set which is the sum of all the sets of the collection G.

For example, any closed set is semi-closed, as is also any totally disconnected set or the sum of any closed set and any set of dimension zero. Any null collection of disjoint closed sets (i.e., a collection having only a finite number of elements of diameter greater than any preassigned $\epsilon > 0$) is semi-closed. The collection of components of any closed set K is semi-closed, as is also this collection together with an arbitrary null collection of disjoint closed sets, no one of which intersects K.

The principal object of the present paper will be to develop conditions under which the complements of semi-closed sets and collections in various continuum spaces will be connected and locally connected.

2. We begin with some results giving fundamental relations between these sets and collections and upper semi-continuous collections.[1]

(2.1) Theorem. *If a collection G of disjoint closed sets is upper semi-continuous, then in order that G be semi-closed it is necessary and sufficient that the decomposition of S into the sets of G and the individual points of $S - G^*$ be upper semi-continuous.*

This theorem follows immediately from the definitions of semi-closed and of upper semi-continuous collections.

(2.11) Corollary. *If G is any upper semi-continuous collection of disjoint closed sets filling up S, and if G_0 is the set of all non-degenerate elements of the collection G, then any subcollection G_1 of G such that $G_0 \subset G_1 \subset G$ is semi-closed.*

(2.2) *The collection G of all components of any semi-closed set K in a compact space S is upper semi-continuous.*

For if this were not so, there would exist a convergent sequence $g_1, g_2, \cdots$ of sets of G such that if $L = \lim (g_i)$, then for some $g \epsilon G$ we have

$$L \cdot g \neq 0 \neq L \cdot (S - g).$$

Received August 24, 1936.

[1] See R. L. Moore, Transactions of the American Mathematical Society, vol. 27 (1925), pp. 416–428. As used in the present paper, a collection G is upper semi-continuous provided that for every convergent sequence of elements (g_i) of G whose limit set L intersects $g \epsilon G$ we have $L \subset g$. For compact spaces this is equivalent to Moore's original definition. See ref. 3.

Since L is connected and g is a component of K, it follows that $L \cdot (S - K) \neq 0$. But then (g_i) would necessarily converge to a single point of $S - K$, contrary to $L \cdot g \neq 0$.

(2.3) THEOREM. *If K is a semi-closed subset of a compact space S, there exists a monotone transformation*[2] $T(S) = S'$ *of S into a compact space S' such that for each $y \in S'$, $T^{-1}(y)$ is either a component of K or a single point of $S - K$. Thus T is a homeomorphism on $S - K$.*

For by (2.2) and (2.1) the decomposition of S into the components of K and the individual points of $S - K$ is upper semi-continuous. Hence by well known results[3] this decomposition is equivalent to a monotone transformation $T(S) = S'$ satisfying all the requirements of (2.3).

(2.4) THEOREM. *In order that a set K be semi-closed it is sufficient, and in case the space is compact it is also necessary, that for each $\epsilon > 0$ the sum of all the components of K of diameter $\geqq \epsilon$ be a closed set.*

Proof. To prove the sufficiency, let $g_1, g_2, \cdots$ be any convergent sequence of components of K whose limit set L intersects $S - K$. Then this sequence must be a null sequence. For if it contains an infinite subsequence (g_{n_i}), each element of which is of diameter greater than some preassigned $\epsilon > 0$, it would follow from our hypothesis that Σg_{n_i} is contained in a closed set which in turn is contained in K, and this is impossible since (g_{n_i}) converges to L and $L \cdot (S - K) \neq 0$. Thus (g_i) is a null sequence and hence L must reduce to a single point of $S - K$.

To prove the necessity, let us suppose S is compact, let $\epsilon > 0$ be given, and let K_ϵ be the sum of all components of K of diameter $\geqq \epsilon$. Then if K is not closed, it readily follows that there exists in K_ϵ a convergent sequence of components $g_1, g_2, \cdots$ of K whose limit set L is not contained in K_ϵ. But this is impossible because $\delta(g_i) \geqq \epsilon$, $i = 1, 2, \cdots$, gives $\delta(L) \geqq \epsilon$. Also L is connected, and since K is semi-closed, we must have $L \subset K$. Thus L is contained in some component g of K which in turn belongs to K_ϵ.

(2.41) COROLLARY. *If S is compact and $K \subset S$ is semi-closed, the sum of all non-degenerate components of K is an F_σ.*

3. We shall develop next certain notions of separation which will be needed in what follows.

If K is any set, a continuum N in K is said to *separate* K provided there exists a separation $K - N = K_1 + K_2$, where both K_1 and K_2 intersect the component of K containing N; N is said to *separate K locally* provided N separates some open subset of K containing N.

(3.1) THEOREM. *If $T(A) = B$ is monotone, where A is a compact continuum,*

[2] A single-valued continuous transformation $T(A) = B$ is said to be *monotone* provided that for each $b \in B$, $T^{-1}(b)$ is a connected set. See C. B. Morrey, American Journal of Mathematics, vol. 57 (1935), pp. 17–50.

[3] See Alexandroff, Mathematische Annalen, vol. 96 (1926), pp. 555–571; Kuratowski, Fundamenta Mathematicae, vol. 11 (1928), pp. 169–185.

a subcontinuum K of B will locally separate B if and only if $T^{-1}(K)$ locally separates A.

Proof. Suppose $T^{-1}(K)$ locally separates A. Then there exists a neighborhood U of $T^{-1}(K)$ in A and a separation

$$U - T^{-1}(K) = U_1 + U_2, \tag{i}$$

where both U_1 and U_2 intersect the component C of U containing $T^{-1}(K)$. Since $T^{-1}(K) \subset U$, there exists a neighborhood V of K such that

$$T^{-1}(V) \subset U. \tag{ii}$$

Since $T^{-1}(V)$ is open in A, by (i) we have a separation

$$T^{-1}(V) - T^{-1}(K) = V_1' + V_2', \tag{iii}$$

where $V_i' = U_i \cdot T^{-1}(V)$ $(i = 1, 2)$ and clearly both V_1' and V_2' intersect the component C' of $T^{-1}(V)$ containing $T^{-1}(K)$. Applying T to (iii), we get

$$V - K = T(V_1') + T(V_2'). \tag{iv}$$

Now since T is monotone, it follows that $T^{-1}T(V_i') = V_i'$ $(i = 1, 2)$, and hence, as A is compact, the sets $T(V_1')$ and $T(V_2')$ are mutually separated. Furthermore, each of these sets intersects the component of V containing K, since both V_1' and V_2' intersect C'. Thus K separates V and accordingly locally separates B.

To prove the converse, let us assume that K locally separates B. Then there exists a neighborhood V of K in B and a separation

$$V - K = V_1 + V_2,$$

where both V_1 and V_2 intersect the component C of V containing K. Now if we set

$$T^{-1}(V) = U, \qquad T^{-1}(V_i) = U_i \qquad (i = 1, 2),$$

U is open in B and we have a separation

$$U - T^{-1}(K) = U_1 + U_2,$$

since T is continuous. Now since T is monotone, it follows that $T^{-1}(C)$ is the component of U containing $T^{-1}(K)$. Thus $T^{-1}(K)$ separates U and locally separates A, since $U_i \cdot T^{-1}(C) \neq 0$ $(i = 1, 2)$.

(3.2) THEOREM. *Let F be a totally disconnected set of non-local-separating-points of a locally connected continuum S such that only a countable number of components of any irreducible cutting of S between two points a and b intersect F. Then $S - F$ is connected and locally connected.*

Proof. Let R be any region in S. We shall prove that $R \cdot (S - F)$ is connected. Clearly our theorem results at once from this. Suppose, on the contrary, that $R \cdot (S - F)$ is separated between two of its points a and b. Then

$F + S - R$ cuts S between a and b. Accordingly, it contains a closed irreducible cutting X of S between a and b. By hypothesis only a countable number of the components of X intersect F.

Now since R is connected and $R \supset a + b$, we have $R \cdot X \neq 0$, and $R \cdot X$ cuts R between a and b. But since $R \cdot X \subset F$, it follows that $R \cdot X$ is countable; since $R \cdot X$ is countable and closed in R, it must contain an isolated point, and this must be a local separating point. This contradicts the hypothesis that no point of F is a local separating point.

Thus $R \cdot (S - F)$ is connected and our theorem follows.

4. We now apply the preceding results to obtain conclusions concerning the connectivity and local connectivity of the complement of a semi-closed set.

(4.1) THEOREM. *Let F be a semi-closed subset of a compact locally connected continuum S such that no component of F separates S locally and such that for any irreducible cutting K of S between two points, $K \cdot F$ is contained either in a countable number of components of K or in a countable number of components of F. Then $S - F$ is connected and locally connected.*

Proof. By (2.3) there exists a monotone transformation $T(S) = W$ such that for each $w \in W$, $T^{-1}(w)$ is either a component of F or a single point of $S - F$. Then W is a locally connected continuum and $T(F)$ is a totally disconnected set of non-local-separating-points of W [by (3.1), since no $T^{-1}(w)$ locally separates S for $w \in T(F)$].

Now let X be any irreducible cutting of W between two points a and b. Then $T^{-1}(X)$ separates S between points $a' \in T^{-1}(a)$ and $b' \in T^{-1}(b)$. Let Y be a subset of $T^{-1}(X)$ separating S irreducibly between a' and b'. Then, by hypothesis, either only a countable number of components $Y_1, Y_2, \cdots$ of Y can intersect F or only a countable number of components of F intersect Y.

Now we must have $T(Y) = X$. For if not, since $T(Y) \subset X$ and X cuts W irreducibly between a and b, there would exist a connected set N such that $a + b \subset N \subset W - T(Y)$, and this is impossible, since from the fact that T is monotone, $T^{-1}(N)$ is connected and contains $a' + b'$ but does not intersect Y.

Thus $T(Y) = X$. Now in either of the two possible cases it is clear that there exists a countable sequence $X_1, X_2, \cdots$ of components of X such that $T(F \cdot Y) \subset X_1 + X_2 + \cdots$. But it follows at once that $T(F \cdot Y) = T(F) \cdot T(Y)$, since $z \in T(F) \cdot T(Y)$ gives $T^{-1}(z) =$ a component of F. Thus $\Sigma X_i \supset T(F) \cdot T(Y) = T(F) \cdot X$. Hence only a countable number of the components of any irreducible cutting of W between two points can intersect $T(F)$.

Thus by (3.2), $W - T(F)$ is connected and locally connected. But by the definition of T we have $T(S - F) = W - T(F)$ and T is a homeomorphism on the set $S - F$. Therefore $S - F$ is connected and locally connected.

(4.2) THEOREM. *Let F be any semi-closed subset of a compact uni-coherent*[4]

[4] For definitions of these terms, see Kuratowski, Fundamenta Mathematicae, vol. 12 (1928), p. 24.

(n-coherent, $\aleph_0$-coherent)[4] *locally connected continuum S such that no component of F separates S (separates S locally). Then $S - F$ is connected and locally connected.*[5]

Proof. By uni-coherence of S and the fact that no component of F separates S, it follows that in any case no component of F separates S locally.

Now if X is any irreducible cutting of S between two points a and b and if S_a and S_b are the components of $S - X$ containing a and b respectively, then, since we can express S as the sum of two continua in the form

$$S = (S_a + X) + (S - S_a),$$

where $(S_a + X)\cdot(S - S_a) = X$, it follows that X can have at most one, n, $\aleph_0$ components according as S is uni-, n-, $\aleph_0$-coherent. Thus surely X can in any case have only a countable number of components intersecting F. Therefore, by (4.1), $S - F$ is connected and locally connected.

(4.3) Theorem. *Let M be a subcontinuum of a compact uni-coherent locally connected continuum S such that the complementary domain boundaries of M in S are disjoint, and form a semi-closed collection, but no one of them separates M. Then if B denotes the sum of these boundaries, $M - B$ is connected and locally connected. Furthermore, $M - B$ is homeomorphic with the complement of a countable set of points on a cyclic monotone image Σ of S, and thus if S is a topological sphere, so is Σ; hence if M is non-dense on S, $M - B$ is homeomorphic with the set of all irrational points in a plane.*

Proof. Let the domain boundaries be $B_1, B_2, \cdots$. Then by hypothesis[6] $B_1, B_2, \cdots$ are the components of B and B is semi-closed. Now for each i, let F_i denote the set B_i plus all complementary domains of B_i except the one containing $M - B_i$. (Note that the set $M - B_i$ is connected by hypothesis.) Let $F = \Sigma F_i$. Then for each i, F_i is a component of F and it readily follows that F is semi-closed. Since, for each i, $S - F_i$ is a single complementary domain of B_i, no set F_i separates S. Therefore, by (4.2), $S - F$ is connected and locally connected.

Now $S - F = M - B$. For let $x \in S - F$. Then $x \in M$. For if not, x lies in a complementary domain D_k of M, and this gives $x \subset D_k \subset D_k + B_k \subset F_k \subset F$, an impossible result. Thus $x \in M$, and since $B \subset F$, we have $x \in M - B$, whence $S - F \subset M - B$. On the other hand, $F \supset B$ and $F \subset B + S - M$ gives $S - F \supset M - B$, whence $S - F = M - B$. Accordingly, $M - B$ is connected and locally connected.

Now by (2.3) there exists a monotone transformation $T(S) = \Sigma$ such that for each $x \in \Sigma$, $T^{-1}(x)$ is either a component of F or a single point of $S - F$, because F is semi-closed. Thus $M - B$ $(= S - F)$ is homeomorphic with $\Sigma - T(F)$. Also, Σ is cyclic, since no set $T^{-1}(x)$, $x \in \Sigma$, separates S.

[5] A result closely related to this theorem has been found by R. L. Wilder. See his abstract in the Bulletin of the American Mathematical Society, vol. 34 (1928), p. 426, no. 22.

[6] It is readily seen that, when added together, the elements of any countable semi-closed collection of disjoint continuua (X_i) in a compact space form a semi-closed set whose components are the X_i.

Now since if M is non-dense in S, F is dense in S and $T(F)$ is dense in Σ, and since $T(F)$ is countable, it follows that in this case $M - B$ is homeomorphic with the complement on Σ of a countable dense set.

If S is a topological sphere, so is Σ, since Σ is the cyclic monotone image[1] of S. Thus if M is non-dense in S, $T(F)$ is countable and dense in Σ, and hence it is isotopic with the set of rational points on a sphere or plane. Thus $M - B$ is homeomorphic with the set of irrational points on a sphere or plane.

As a simple application, let S be a sphere and let M be a locally connected continuum on S having no local separating point. The conditions of (4.3) are then satisfied and accordingly $M - B$ is homeomorphic with the complement of a countable set of points on a sphere, and if M is non-dense, $M - B$ is homeomorphic with the set of all irrational points on a plane.

THE UNIVERSITY OF VIRGINIA.

CRITERIA FOR THE COMPOSITENESS OF FINITE GROUPS

By Louis Weisner

1. **Introduction.** Tchounikhin has recently established a criterion for the compositeness of a finite group[1] of which the following is a modification.

Theorem 1. *Let G be a group of order $g = p^\alpha m$, where p is a prime that does not divide m, and let P be a subgroup of order p^α of G. G has an invariant subgroup of index p that does not include a particular element S of P if and only if P has a maximal subgroup P_1 which does not include S, such that every conjugate of S^l (l any integer) under G that is contained in P has the form S^lV, where V is an element of P_1.*

I propose to show in the present paper that the condition of the theorem is satisfied under fairly general assumptions concerning the relation of P to G, thus deducing new compositeness criteria.

2. **Notation.** The notations of Theorem 1 for G and P will be employed throughout this paper. The normalizer in G of an element A or subgroup A of G will be denoted by $N(A)$. The cross-cut of two groups Γ_1 and Γ_2 will be denoted by (Γ_1, Γ_2) and the group they generate by $\{\Gamma_1, \Gamma_2\}$.

3. **Proof of Theorem 1.** Let G' be an invariant subgroup of index p in G which does not include S. $P_1 = (P, G')$ is clearly a maximal subgroup of P and a Sylow subgroup of G'. The commutator of S^l and any element of G is an element of G'; hence, if this commutator is an element of P, it is an element of P_1. The condition of the theorem is therefore necessary.

In proving the condition sufficient we shall suppose, without introducing any change of notation, that G is a regular permutation group on the symbols $x_1, \cdots, x_g$. By a suitable choice of the notation we may suppose that

$$y_0 = x_1 + \cdots + x_n \qquad (n = p^{\alpha-1})$$

belongs to P_1. The conjugates of y_0 under G are linearly independent, since no two of them have a term in common. The permutations of P transform y_0 into

$$y_k = S^k y_0 \qquad (k = 0, 1, \cdots, p-1),$$

each of which is an invariant of P_1, since P_1 is an invariant subgroup of P. The function

$$\varphi_1 = \sum_{k=0}^{p-1} \epsilon^{-k} y_k,$$

Received June 3, 1936; presented to the American Mathematical Society, October 26, 1935.

[1] Serge Tchounikhin, *Über einige Sätze der Gruppentheorie*, Mathematische Annalen, vol. 112 (1935), p. 92.

where ϵ is a primitive p-th root of unity, is therefore an absolute invariant of P_1. It is, however, a relative invariant of P, since

$$S^c\varphi_1 = \epsilon^c\varphi_1 \qquad (c = 0, 1, \cdots, p-1).$$

Moreover, since the only permutations of G that permute $y_0, \cdots, y_{p-1}$ among themselves are those of P, every permutation of G that transforms φ_1 into a numerical multiple of itself is contained in P.

Let

$$G = PT_1 + \cdots + PT_m \qquad (T_1 = 1)$$

be a decomposition of G into cosets as regards P; and let

$$\varphi_i = T_i\varphi_1 \qquad (i = 1, \cdots, m).$$

These functions are distinct. If B is a permutation of G and T_iB is in the j-th coset, $T_iB = UT_j$, where U is a permutation of P. Therefore[2]

$$B\varphi_i = (T_iB)\varphi_1 = (UT_j)\varphi_1 = T_j(\lambda\varphi_1) = \lambda\varphi_j \qquad (\lambda^p = 1).$$

Every permutation of G therefore transforms each of the functions $\varphi_1, \cdots, \varphi_m$ into one of these functions multiplied by a power of ϵ.[3] Their product

$$\phi = \varphi_1 \cdots \varphi_m$$

is therefore an invariant of G. We proceed to prove that ϕ is a *relative* invariant of G.

Suppose that, for a certain index i,

$$(1) \qquad S\varphi_i = \epsilon^c\varphi_i \qquad (0 \leqq c \leqq p-1).$$

Then

$$(T_iS)\varphi_1 = \epsilon^c(T_i\varphi_1),$$

$$(2) \qquad (T_iST_i^{-1})\varphi_1 = \epsilon^c\varphi_1.$$

We have seen that the only permutations of G that transform φ_1 into a numerical multiple of itself are those of P; therefore $T_iST_i^{-1}$ is contained in P. By hypothesis $T_iST_i^{-1} = SV$, where V is a permutation of P_1. Therefore, since $V\varphi_1 = \varphi_1$,

$$(3) \qquad (T_iST_i^{-1})\varphi_1 = (SV)\varphi_1 = \epsilon\varphi_1.$$

Comparing (2) with (3), we infer that (1) is possible only if $c = 1$. Supposing the notation chosen so that

[2] In the equations which follow the permutations operate from left to right.

[3] A representation of G as a monomial group therefore arises, and the proof of the theorem may be completed by employing the theory of monomial groups, following W. Burnside, *Theory of Groups*, second edition, 1911, p. 325. This is the method employed by Tchounikhin, loc. cit. I prefer, however, to use the more elementary concept of relative invariant of a permutation group, following H. F. Blichfeldt, *Theorems on simple groups*, Transactions of the American Mathematical Society, vol. 11 (1910), pp. 1–14.

$$S\varphi_i = \epsilon\varphi_i \qquad (i = 1, \cdots, r),$$
$$S\varphi_i \neq \epsilon\varphi_i \qquad (i = r + 1, \cdots, m),$$

we have

$$S(\varphi_1 \cdots \varphi_r) = \epsilon^r(\varphi_1 \cdots \varphi_r),$$

where $r \geqq 1$, since $S\varphi_1 = \epsilon\varphi_1$. If $r = m$, ϕ is a relative invariant of G, since $(m, p) = 1$.

If $r < m$, each of the functions $\varphi_{r+1}, \cdots, \varphi_m$ is transformed by S into *another* of the set, multiplied by a power of ϵ. Aside from these multipliers, S permutes these functions according to a permutation whose order is a power of p. Their number $m - r$ is therefore a multiple of p, so that $(r, p) = 1$ in any case. If $(\varphi_{r+1}, \cdots, \varphi_{r+p^e})$ is a cycle of the permutation in question,

$$S\varphi_{r+1} = \epsilon_1\varphi_{r+2}, \qquad S\varphi_{r+2} = \epsilon_2\varphi_{r+3}, \cdots, S\varphi_{r+p^e} = \epsilon_{p^e}\varphi_{r+1},$$

where $\epsilon_1, \epsilon_2, \cdots, \epsilon_{p^e}$ are powers of ϵ. Denoting their product by

$$\epsilon^c \qquad (0 \leqq c \leqq p - 1),$$

we have

$$S^{p^e}\varphi_{r+1} = \epsilon^c\varphi_{r+1}.$$

Hence

$$(T_{r+1}S^{p^e}T_{r+1}^{-1})\varphi_1 = \epsilon^c\varphi_1.$$

$T_{r+1}S^{p^e}T_{r+1}^{-1}$ is therefore a permutation of P and consequently, by hypothesis, is a permutation of P_1, since S^{p^e} is a permutation of P_1. It follows that $c = 0$, so that

$$S(\varphi_{r+1} \cdots \varphi_{r+p^e}) = \varphi_{r+1} \cdots \varphi_{r+p^e}.$$

Treating the remaining functions in the same way we conclude that $S\phi = \epsilon^r\phi$. Therefore, since $(r, p) = 1$, ϕ is a relative invariant of G. Those permutations of G which transform ϕ into itself form an invariant subgroup of G of index p which does not include S. The proof of the theorem is complete.

4. **Theorems involving primitive elements.** An element of a group is *imprimitive* or *primitive* according as it is or is not contained in every maximal subgroup of the group. The appropriateness of these terms may be realized by comparison with the concept of a primitive element of a field, considering that a primitive element of a group is included in at least one set of independent generators of the group, while an imprimitive element is not.[4]

THEOREM 2. *If a primitive element of S of P is commutative with every element of G whose order is prime to p, G has an invariant subgroup of index p that does not include S.*

[4] See Miller, Blichfeldt and Dickson, *Finite Groups*, 1916, p. 71.

This theorem is an immediate consequence of Theorem 1, as the conditions of Theorem 1 are fulfilled by S and any maximal subgroup P_1 of P that does not include S.

Repeated applications of Theorem 2 yield the following theorem: *If every element of G whose order is prime to p is commutative with each element whose order is a power of p, then G is the direct product of P and a group of order m.*[5]

If, in addition to the assumptions of Theorem 2, we assume that P is invariant under G, we may conclude that a maximal subgroup of P is invariant under G. For if G' is the invariant subgroup of G, whose existence is asserted by Theorem 2, then (P, G') is a maximal subgroup of P which is invariant under G.

THEOREM 3. *If P is invariant under G and a primitive element of P is commutative with every element of G whose order is prime to p, then P has a maximal subgroup which is invariant under G.*

In particular, *an automorphism of P, of order prime to p, which is commutative with a primitive element of P is commutative with some maximal subgroup of P.*

THEOREM 4. *If a primitive element of P is invariant under $N(P)$ and under every Sylow subgroup of G in which it enters, then G has an invariant subgroup of index p.*

Let S be the primitive element in question. Since S is invariant under every Sylow subgroup of G in which it enters, the same is true of every conjugate of S. Let S_1 be a conjugate of S under G, and suppose that S_1 is contained in P. Since S and S_1 are conjugates under G, they are conjugates under $N(P)$.[6] But S is invariant under $N(P)$; hence $S = S_1$. Therefore no conjugate of S under G, except S itself, is contained in P. The same being true of every power of S, the conditions of Theorem 1 are fulfilled by S and any maximal subgroup P_1 of P that does not include S. Therefore G has an invariant subgroup of index p that does not include S.

5. **General theorems.** When the distribution of the elements of P into conjugate sets with respect to G is known, all invariant subgroups of G of index p may be determined by Theorem 1, if any exist. In cases where the information concerning this distribution is inadequate, but sufficient information is available concerning the distribution into conjugate sets of the elements of a subgroup Γ of G that includes P, the following theorem may be found useful.

THEOREM 5. *Let S be an element of P and Γ a subgroup of G that includes P. If every two conjugates of S^l (l any integer) under G that are contained in Γ are conjugates under Γ, and if Γ has an invariant subgroup of index p that does not include S, the same is true of G.*[7]

[5] Burnside, *Theory of Groups*, p. 327, Corollary 1.

[6] Burnside, *Theory of Groups*, p. 155.

[7] Proved by W. K. Turkin, *Ein neues Kriterium der Einfachheit einer endlichen Gruppe*, Mathematische Annalen, vol. 111 (1935), p. 281, subject to the assumption that the order of Γ and the index of Γ in G are relatively prime. Other criteria of this type are given by G. Frobenius, *Ueber auflösbare Gruppen*, III and V, Sitzungsberichte Berlin, 1901, pp. 865 and 1324.

If Γ' is the invariant subgroup of index p of Γ, $P_1 = (P, \Gamma')$ does not contain S and is a maximal subgroup of P. By Theorem 1, applied to Γ, every conjugate of S^l under Γ that is contained in P has the form S^lV, where V is an element of P_1. All conjugates of S^l under G that are elements of P are accounted for, since two conjugates of S^l under G that are contained in P are conjugates under Γ. The theorem now follows from Theorem 1, applied to G.

To apply this theorem effectively, it is desirable to know under what circumstances all the conjugates under G of an element of Γ, that are contained in Γ, are conjugates under Γ. The next theorem will be found useful in this connection.

THEOREM 6. *Let Γ be a subgroup of G, and C an element or subgroup of Γ. All the conjugates of C under G which are contained in Γ are conjugates under $N(\Gamma)$ if and only if every two conjugates of Γ under G that contain C are transformable into one another by an element of G that is commutative with C.*

To prove the condition necessary, we suppose that C is contained in two conjugates Γ_1 and Γ_2 of Γ. We do not assume that Γ, Γ_1 and Γ_2 are distinct. If A is an element of G that transforms Γ_1 into Γ_2 and A transforms C into C_0, C and C_0 are contained in Γ_2 and are conjugates under $N(\Gamma_2)$ since, by hypothesis, all the conjugates of C under G that are contained in Γ are conjugates under $N(\Gamma)$. An element B of $N(\Gamma_2)$ therefore exists which transforms C into C_0. We now have

$$A^{-1}\Gamma_1A = \Gamma_2, \qquad A^{-1}CA = C_0, \qquad B^{-1}\Gamma_2B = \Gamma_2, \qquad B^{-1}CB = C_0.$$

It follows that AB^{-1} is commutative with C and transforms Γ_1 into Γ_2.

To prove the condition sufficient, let C_1 and C_2 be two conjugates of C under G that are contained in Γ. Let T be an element of G that transforms C_1 into C_2, and suppose that T transforms Γ into Γ_0. Since C_2 is contained in Γ and in Γ_0, an element U of G exists which is commutative with C_2 and transforms Γ into Γ_0. Therefore

$$T^{-1}\Gamma T = \Gamma_0, \qquad T^{-1}C_1T = C_2, \qquad U^{-1}\Gamma U = \Gamma_0, \qquad U^{-1}C_2U = C_2.$$

It follows that TU^{-1} is an element of $N(\Gamma)$ that transforms C_1 into C_2.

It will be noticed that this theorem provides a condition that the conjugates of C under G which are contained in Γ be conjugates under $N(\Gamma)$ rather than under Γ. It is, however, possible to choose Γ so that every two elements or subgroups of Γ which are conjugates under $N(\Gamma)$ are conjugates under Γ. This will surely be the case if $\Gamma = N(\Gamma)$, that is, if Γ is its own normalizer in G, or if every element of $N(\Gamma)$ which is not contained in Γ is commutative with each element of Γ.

6. **Case in which P is abelian.** If G' is an invariant subgroup of G of index p, $(N(P),G')$ is an invariant subgroup of $N(P)$ of index p. Therefore, if G has an invariant subgroup of index p, the same is true of $N(P)$.

The converse is true when P is an abelian group. For in this case two ele-

ments of P which are conjugates under G are conjugates under $N(P)$. It follows from Theorem 5, with $\Gamma = N(P)$, that if $N(P)$ has an invariant subgroup of index p, the same is true of G.

THEOREM 7. *If P is an abelian group, G has an invariant subgroup of index p if and only if $N(P)$ has an invariant subgroup of index p.*

If G has an invariant subgroup of index p, and P' is a subgroup of order $p^{\alpha-1}$ of the invariant subgroup of $N(P)$ of index p whose existence we have proved, P' is an invariant subgroup of $N(P)$, and the corresponding quotient group has only one subgroup of order p, this subgroup being an invariant Sylow subgroup. Since, on the other hand, $N(P)/P'$ has an invariant subgroup of index p, an element of order p of $N(P)/P'$ is invariant under this group. Therefore, if G has an invariant subgroup of index p, P has a maximal subgroup P' which is invariant under $N(P)$, such that $N(P)/P'$ includes an invariant element of order p.

The converse is true when P is abelian. For if a maximal subgroup P' of P is invariant under $N(P)$, and $N(P)/P'$ has an invariant element of order p, $N(P)$ has an invariant subgroup of index p. It follows from Theorem 7 that G has an invariant subgroup of index p.

THEOREM 8. *If P is an abelian group, G has an invariant subgroup of index p if and only if P has a maximal subgroup P' which is invariant under $N(P)$, such that $N(P)/P'$ includes an invariant element of order p.*

If $(g, p - 1) = 1$, an element of order p of $N(P)/P'$ cannot be a conjugate of any of its powers, except the first power, and is therefore invariant under $N(P)/P'$. We therefore have the following

THEOREM 9. *If P is an abelian group, if $(g, p - 1) = 1$, and if a maximal subgroup of P is invariant under $N(P)$, then G has an invariant subgroup of index p.*

Let $p^{n_1}, p^{n_2}, \cdots, p^{n_r}$ be the invariants of the abelian group P, arranged in descending order of magnitude. Those elements of P whose orders divide p^{n_1-1} form a characteristic subgroup of P whose order is $p^{\alpha-1}$ if the largest invariant of P is unrepeated; that is, if $n_1 > n_2$. From the preceding theorem we now have

THEOREM 10. *If P is an abelian group of type $(n_1, n_2. \cdots, n_r)$, if $n_1 > n_2 \geqq n_3 \geqq \cdots \geqq n_r$, and if $(g, p - 1) = 1$, then G has an invariant subgroup of index p.*[8]

7. $\Gamma = N(L)$, **where L is a subgroup of the central of P.** We proceed to prove

THEOREM 11. *If a subgroup L of the central of P is invariant under $N(P)$ and under every subgroup of order p^α of G that contains L, and if $N(L)$ has an invariant subgroup of index p, then G has an invariant subgroup of index p.*

Since L is invariant under every Sylow subgroup of G into which it enters,

[8] If $r = 1$ or 2, G has an invariant subgroup of index p^α. Burnside, *Theory of Groups*, p. 327, Corollary 2.

every conjugate of L has the same property. If a conjugate L_1 of L were contained in P, L and L_1 would be conjugates under $N(P)$, since they are invariant subgroups of P. But L is invariant under $N(P)$; hence $L = L_1$. Therefore, *two distinct conjugates of L cannot be contained in the same Sylow subgroup of G.* Again, since L is a subgroup of the central of P, and every conjugate of P under G contains a conjugate of L, L is a subgroup of the central of every Sylow subgroup of G into which it enters. The only subgroups of order p^α of G that contain L are those of $N(L)$. If U is an element of G that transforms $N(L)$ into itself, $U^{-1}LU$ is invariant under every subgroup of order p^α of $N(L)$; hence $U^{-1}LU = L$. It follows that $N(L)$ is its own normalizer in G.

Every conjugate of $N(L)$ under G is the normalizer of some conjugate of L. Suppose that an element T, of order a power of p, is contained in $N(L)$ and in $N(L_1)$, where L_1 is a conjugate of L. Since T is commutative with every element of L and L_1, L and L_1 are subgroups of $N(T)$. Two Sylow subgroups J and J_1 of $N(T)$ that contain L and L_1 respectively are conjugates under $N(T)$. If A is an element of $N(T)$ that transforms J into J_1, A must transform L into L_1 and $N(L)$ into $N(L_1)$; otherwise, two distinct conjugates of L would be contained in the same Sylow subgroup of G. It follows from Theorem 6 that two elements of $N(L)$, of order a power of p, that are conjugates under G are conjugates under $N(L)$. By hypothesis $N(L)$ has an invariant subgroup of index p. If S is an element of P that is not contained in this invariant subgroup, the conditions of Theorem 5 are fulfilled by S and $\Gamma = N(L)$. Therefore G has an invariant subgroup of index p that does not include S.

HUNTER COLLEGE OF THE CITY OF NEW YORK.

TRIPLES OF CONJUGATE HARMONIC FUNCTIONS AND MINIMAL SURFACES

By J. W. Hahn and E. F. Beckenbach

A surface S is said to be given in terms of isothermic parameters u, v, provided the representation

$$S: x_j = x_j(u, v), \qquad j = 1, 2, 3, \qquad (u, v) \text{ in } D, \tag{1}$$

where D is some finite domain of definition, is such that

$$E = G = \lambda(u, v), \qquad F = 0, \tag{2}$$

where

$$E = x_{1,u}^2 + x_{2,u}^2 + x_{3,u}^2, \qquad F = x_{1,u}x_{1,v} + x_{2,u}x_{2,v} + x_{3,u}x_{3,v},$$
$$G = x_{1,v}^2 + x_{2,v}^2 + x_{3,v}^2,$$

the second subscripts denoting differentiation. Such a representation is conformal except where $\lambda(u, v) = 0$.

A theorem of Weierstrass states that a necessary and sufficient condition that a surface S, given in terms of isothermic parameters, be minimal is that the coördinate functions be harmonic. Then in any simply connected part of D, the functions x_j are the real parts of analytic functions,

$$x_j = \Re f_j(w), \qquad w = u + iv,$$

and (2) is equivalent to

$$\sum_{j=1}^{3} f_j'^2(w) = 0. \tag{3}$$

If an isothermic representation (1) of the minimal surface S is such that one of the coördinate functions is identically zero, say $x_3(u, v) \equiv 0$, then either $x_1(u, v) + ix_2(u, v)$ or $x_2(u, v) + ix_1(u, v)$ is an analytic function of the complex variable $w = u + iv$, and $x_1(u, v)$ and $x_2(u, v)$ are said to form a *couple of conjugate harmonic functions*. By analogy, the coördinate functions of any minimal surface in isothermic representation have been called a *triple of conjugate harmonic functions*.[1]

The analogy here indicated between analytic functions of a complex variable and isothermic representations of minimal surfaces has often been noted, and since the time of Weierstrass has served as a guiding principle in the study of minimal surfaces. It is the purpose of the present paper to pursue this analogy from the coefficients viewpoint.

Received March 30, 1936.

[1] E. F. Beckenbach and T. Radó, *Subharmonic functions and minimal surfaces*, Trans. Amer. Math. Soc., vol. 35 (1933), pp. 648–661.

If $x_j(u, v)$ is harmonic for (u, v) in the domain D, and if P is an interior point of D, then $x_j(u, v)$ can be represented in the neighborhood of P by a Fourier series,

$$x_j = a_j + \sum_{k=1}^{\infty} r^k(a_{j,k} \cos k\theta + b_{j,k} \sin k\theta), \tag{4}$$

where r and θ are polar coördinates with pole at P. Two such functions $x_1(u, v)$ and $x_2(u, v)$ form a couple of conjugate harmonic functions if and only if

$$a_{1,k} = \pm b_{2,k}, \qquad a_{2,k} = \mp b_{1,k} \qquad k = 1, 2, 3, \cdots. \tag{5}$$

In Lemma 1 we shall determine necessary and sufficient conditions, analogous to (5), on the coefficients in (4), in order that three such functions $x_1(u, v)$, $x_2(u, v)$, $x_3(u, v)$ shall form a triple of conjugate harmonic functions. The discussion, while restricted to three functions, holds equally well for sets of n conjugate harmonic functions, as coördinate functions in isothermic representations of minimal surfaces in euclidean n-space.

LEMMA 1. *In order that the harmonic functions*

$$x_j = a_j + \sum_{k=1}^{\infty} r^k(a_{j,k} \cos k\theta + b_{j,k} \sin k\theta) \qquad j = 1, 2, 3,$$

form a triple of conjugate harmonic functions, it is necessary and sufficient that

$$\sum_{l=1}^{k-1} l(k-l) \sum_{j=1}^{3} (a_{j,l}a_{j,k-l} - b_{j,l}b_{j,k-l}) = 0 \qquad k = 2, 3, 4, \cdots, \tag{6}$$

$$\sum_{l=1}^{k-1} l(k-l) \sum_{j=1}^{3} (a_{j,l}b_{j,k-l} + b_{j,l}a_{j,k-l}) = 0 \qquad k = 2, 3, 4, \cdots. \tag{7}$$

Proof. We have

$$f_j(w) = a_j - ib_j + \sum_{k=1}^{\infty} (a_{j,k} - ib_{j,k})w^k, \qquad w = r(\cos\theta + i \sin\theta),$$

whence

$$\sum_{j=1}^{3} f_j'^2(w) = \sum_{j=1}^{3} \sum_{k=2}^{\infty} \sum_{l=1}^{k-1} l(k-l)(a_{j,l} - ib_{j,l})(a_{j,k-l} - ib_{j,k-l})w^{k-2}. \tag{8}$$

That (6) and (7) are equivalent to (3) follows from (8).

LEMMA 2. *Let S and S' be minimal surfaces given in isothermic representation respectively by*

$$x_j = a_j + \sum_{k=1}^{\infty} r^k(a_{j,k} \cos k\theta + b_{j,k} \sin k\theta) \qquad (j = 1, 2, 3), \tag{9}$$

$$y_j = A_j + \sum_{k=1}^{\infty} r^k(A_{j,k} \cos k\theta + B_{j,k} \sin k\theta) \qquad (j = 1, 2, 3)$$

in the neighborhood of $r = 0$. *If* $x_i = y_i$, $j = 1, 2, 3$, *on a set* R *of points having* $r = 0$ *as a limit point, and if* t *is the first index for which* $\sum_{j=1}^{3} (a_{j,t}^2 + b_{j,t}^2) \neq 0$, *then* t *is the first index for which* $\sum_{j=1}^{3} (A_{j,t}^2 + B_{j,t}^2) \neq 0$.

Proof. Suppose the first index h for which $\sum_{j=1}^{3} (A_{j,h}^2 + B_{j,h}^2) \neq 0$ is greater than t. Then

$$(x_j - y_j)^2 = r^{2t}[a_{j,t}^2 \cos^2 t\theta + b_{j,t}^2 \sin^2 t\theta + 2a_{j,t}b_{j,t} \cos t\theta \sin t\theta + O(r)],$$

where $O(r)$ denotes a quantity $\varphi(r)$ such that $|\varphi(r)/r|$ is bounded for all sufficiently small r. Noting in (6) and (7) that, for our particular t,[2]

$$\sum_{j=1}^{3} a_{j,t}^2 = \sum_{j=1}^{3} b_{j,t}^2 \neq 0, \qquad \sum_{j=1}^{3} a_{j,t}b_{j,t} = 0, \tag{10}$$

we have

$$\sum_{j=1}^{3} (x_j - y_j)^2 = r^{2t}\left[\sum_{j=1}^{3} a_{j,t}^2 + O(r)\right]. \tag{11}$$

Since $\sum_{j=1}^{3} a_{j,t}^2 > 0$, there is a circle about $r = 0$ in which $\sum_{j=1}^{3} (x_i - y_i)^2$ vanishes only at $r = 0$. This is a contradiction of the hypothesis that $r = 0$ is a limit point of zeros of $\sum_{j=1}^{3} (x_j - y_j)^2$. The same reasoning shows that h is not less than t.

In particular, $y_j(u, v) \equiv c_j$, $j = 1, 2, 3$, where each c_j is a constant, defines a minimal surface, so that we have the following result.

THEOREM 1. *Let* S *be a minimal surface given in isothermic representation by* (1). *If* $x_j(u, v) = c_j$, $j = 1, 2, 3$, *where each* c_j *is a constant, on a set* R *of points with a limit point* P *interior to* D, *then* $x_j(u, v) \equiv c_j$ *in* D; *that is,* S *is a point.*

The following familiar theorem is included for the sake of completeness.

THEOREM 2. *Let* S *be a minimal surface given in isothermic representation by* (1). *If* $\lambda(u, v) = 0$ *on a set* R *of points with a limit point* P *interior to* D, *then* $x_j(u, v) \equiv$ *constant in* D; *that is,* S *is a point.*

Proof. Assuming that S is not a point, and taking P as a pole of polar coördinates, so that S is represented near P by (9), we have the same sort of proof as for Theorem 1, except that this time we are concerned not with the function (11) but with the function

[2] One might suspect that the conditions (6) and (7) are reducible to

$$\sum_{j=1}^{3} a_{j,k}a_{j,l} = \sum_{j=1}^{3} b_{j,k}b_{j,l}, \qquad \sum_{j=1}^{3} a_{j,k}b_{j,l} + \sum_{j=1}^{3} a_{j,l}b_{j,k} = 0$$

for all k, l. But for the minimal surface of Enneper, in its standard representation, this is not true, for instance, for $k = l = 2$.

$$\lambda(u, v) = t^2 r^{2t-2}\left[\sum_{j=1}^{3} a_{j,t}^2 + O(r)\right],$$

t being the first index for which $\sum_{j=1}^{3} (a_{j,t}^2 + b_{j,t}^2) \neq 0$.

From Theorem 2 it would follow that unless S is a point, the normal to S exists except at most at the isolated points P where $\lambda = 0$. However, if S is not a point, the formulas for the direction cosines of the normal to S,

$$(12) \qquad X_p = \frac{x_{q,u}x_{s,v} - x_{s,u}x_{q,v}}{\lambda} \qquad p, q, s = 1, 2, 3 \textit{ in cyclic order},$$

which hold except where $\lambda = 0$, reduce by (10) to

$$(13) \qquad X_p = \frac{a_{q,t}b_{s,t} - a_{s,t}b_{q,t}}{\sum_{j=1}^{3} a_{j,t}^2} + O(r);$$

then X_p remains continuous at $r = 0$, insuring the existence of the normal to S even at points where $\lambda = 0$.

THEOREM 3. *Let S be a minimal surface given in isothermic representation by* (1), *and let the direction cosines of the normal to the surface at the image of (u, v) be denoted by $X_j(u, v)$. If $X_j(u, v) = c_j$, $j = 1, 2, 3$, where each c_j is a constant, on a set R of points with a limit point P interior to D, then $X_j(u, v) \equiv c_j$ in D; that is, S is a plane surface.*

Proof.[3] Make a transformation of coördinate axes in the $(x_1x_2x_3)$-space so that the origin is at the image of P and the positive x_3-axis coincides with the positive normal at that point. We have then

$$(14) \qquad X_1 = X_2 = 0, \qquad X_3 = 1$$

at P, and therefore by hypothesis we have (14) on R.

For points of R at which $\lambda \neq 0$, (12) and (14) imply

$$(15) \qquad x_{3,u} = x_{3,v} = 0,$$

while for points at which $\lambda = 0$, (2) implies (15), so that (15) holds at all points of R. Let $f_3(w)$ be a function analytic at P such that

$$x_3(u, v) = \Re f_3(w);$$

from

$$f_3'(w) \equiv x_{3,u} - ix_{3,v}$$

it follows that $f_3'(w)$ vanishes on R and therefore $f_3'(w) \equiv 0$. Then $x_3(u, v) \equiv 0$, and the theorem follows.

[3] Theorems 3 and 4 could be proved also by a consideration of the stereographic projections of the spherical images of the surfaces involved.

THEOREM 4. *Let S and S' be minimal surfaces given in isothermic representation respectively by*

$$x_j = x_j(u, v), \qquad y_j = y_j(u, v) \qquad j = 1, 2, 3, \ (u, v) \text{ in } D$$

and let the direction cosines of the normals to the surfaces at the images of (u, v) be denoted respectively by $X_j(u, v)$, $Y_j(u, v)$; if $x_j(u, v) = y_j(u, v)$, $j = 1, 2, 3$, on a set R of points having a limit point P interior to D, and if $X_j(u, v) = Y_j(u, v)$, $j = 1, 2, 3$, on a set R' of points with the same limit point P, then $x_j(u, v) \equiv y_j(u, v)$ in D; that is, S and S' are coincident surfaces.

Proof. Assume the contrary, that $\sum_{j=1}^{3} (x_j - y_j)^2 \not\equiv 0$. We shall show that our hypotheses are inconsistent, namely that if $\sum_{j=1}^{3} (x_j - y_j)^2 = 0$ on R, but $\sum_{j=1}^{3} (x_j - y_j)^2 \not\equiv 0$, then there is a neighborhood of P in which, except at most at P, $\sum_{j=1}^{3} (X_j - Y_j)^2 \neq 0$, contrary to hypothesis.

If $\sum_{j=1}^{3} (X_j - Y_j)^2 \neq 0$ at P, the above contradiction of hypothesis is trivial, so we take $\sum_{j=1}^{3} (X_j - Y_j)^2 = 0$ at P.

Transform the coördinate axes in the $(x_1x_2x_3)$-space as in the proof of Theorem 3, and do similarly in the $(y_1y_2y_3)$-space, preserving $\sum_{j=1}^{3} (x_j - y_j)^2 = 0$ on R, and take the point P as pole of polar coördinates:

$$x_j = \sum_{k=1}^{\infty} r^k(a_{j,k} \cos k\theta + b_{j,k} \sin k\theta) \qquad (j = 1, 2, 3), \tag{16}$$

$$y_j = \sum_{k=1}^{\infty} r^k(A_{j,k} \cos k\theta + B_{j,k} \sin k\theta) \qquad (j = 1, 2, 3). \tag{17}$$

In the next several paragraphs, a–c, we derive certain relations between the coefficients in (16) and (17).

a. By Lemma 2, if t is the first index for which $\sum_{j=1}^{3} (a_{j,t}^2 + b_{j,t}^2) \neq 0$, then t is the first index for which $\sum_{j=1}^{3} (A_{j,t}^2 + B_{j,t}^2) \neq 0$.

b. Since (14) holds at $r = 0$, (13) gives

$$b_{2,t}a_{3,t} - a_{2,t}b_{3,t} = 0, \tag{18}$$

$$b_{1,t}a_{3,t} - a_{1,t}b_{3,t} = 0, \tag{19}$$

$$a_{1,t}b_{2,t} - a_{2,t}b_{1,t} = \sum_{j=1}^{3} a_{j,t}^2 \neq 0. \tag{20}$$

Equations (18) and (19) are linear and homogeneous in $a_{3,t}$ and $b_{3,t}$; since, by (20), their determinant of coefficients is not zero, it follows that

$$a_{3,t} = b_{3,t} = 0. \tag{21}$$

Equations (10) become then

$$a_{1,t}^2 + a_{2,t}^2 = b_{1,t}^2 + b_{2,t}^2 \neq 0, \tag{22}$$

$$a_{1,t}b_{1,t} + a_{2,t}b_{2,t} = 0. \tag{23}$$

Now (20), (22) and (23) yield

$$b_{1,t} = -a_{2,t}, \qquad b_{2,t} = a_{1,t}. \tag{24}$$

Similarly, we have

$$B_{1,t} = -A_{2,t}, \qquad B_{2,t} = A_{1,t}, \qquad B_{3,t} = A_{3,t} = 0. \tag{25}$$

Making use of (21), (24) and (25), we get

$$\sum_{j=1}^{3} (x_j - y_j)^2 = r^{2t}\left[\sum_{j=1}^{2} (a_{j,t} - A_{j,t})^2 + O(r)\right]. \tag{26}$$

Since $\sum_{j=1}^{3} (x_j - y_j)^2 = 0$ on R, we obtain from (24), (25) and (26),

$$a_{j,t} = A_{j,t}, \qquad b_{j,t} = B_{j,t} \qquad (j = 1, 2).$$

c. Let m be the first index for which

$$\sum_{j=1}^{3} [a_{j,m} - A_{j,m})^2 + (b_{j,m} - B_{j,m})^2] \neq 0; \tag{27}$$

by b, $m > t$. Note that, for $k = m + t$, every term in (6) and (7) which involves an index greater than m involves also an index less than t, and therefore vanishes. This is true also of the corresponding equations involving the A's and B's. Subtracting this second pair of equations from the first pair respectively, and noting that, for $k < m$, $a_{j,k} = A_{j,k}$, $b_{j,k} = B_{j,k}$, we obtain equations which reduce by b to

$$\begin{aligned} &a_{1,t}[(a_{1,m} - A_{1,m}) - (b_{2,m} - B_{2,m})] \\ &\qquad + a_{2,t}[(a_{2,m} - A_{2,m}) + (b_{1,m} - B_{1,m})] = 0, \end{aligned} \tag{28}$$

$$\begin{aligned} &- a_{2,t}[(a_{1,m} - A_{1,m}) - (b_{2,m} - B_{2,m})] \\ &\qquad + a_{1,t}[(a_{2,m} - A_{2,m}) + (b_{1,m} - B_{1,m})] = 0. \end{aligned} \tag{29}$$

Since (22) holds, (28) and (29) yield

$$b_{1,m} - B_{1,m} = -(a_{2,m} - A_{2,m}), \tag{30}$$

$$b_{2,m} - B_{2,m} = a_{1,m} - A_{1,m}. \tag{31}$$

A computation, simplified by (30) and (31), gives

$$\sum_{j=1}^{2} (x_j - y_j)^2 = r^{2m}\left[\sum_{j=1}^{2} (a_{j,m} - A_{j,m})^2 + O(r)\right].$$

Another application of the argument used previously shows, therefore, that

$$a_{j,m} = A_{j,m}, \qquad b_{j,m} = B_{j,m} \qquad (j = 1, 2);$$

therefore, by (27),

$$(a_{3,m} - A_{3,m})^2 + (b_{3,m} - B_{3,m})^2 \neq 0. \tag{32}$$

d. We have

$$x_{j,u} = \sum_{k=t-1}^{\infty} (k+1)\,(a_{j,k+1} \cos k\theta + b_{j,k+1} \sin k\theta)\, r^k,$$

$$x_{j,v} = \sum_{k=t-1}^{\infty} (k+1)\,(b_{j,k+1} \cos k\theta - a_{j,k+1} \sin k\theta)\, r^k,$$

and similar expressions for $y_{j,u}$, $y_{j,v}$. By (12) and the choice of m,

$$X_p - Y_p = \frac{mr^{m-t}}{t \sum_{j=1}^{2} a_{j,t}^2} [M_p - N_p + O(r)],$$

where

$$M_p = (a_{q,t}\, b_{s,m} - b_{q,t}\, a_{s,m} + b_{s,t}\, a_{q,m} - a_{s,t}\, b_{q,m}) \cos (m-t)\,\theta$$
$$+ (- a_{q,t}\, a_{s,m} - b_{q,t}\, b_{s,m} + a_{s,t}\, a_{q,m} + b_{s,t}\, b_{q,m}) \sin (m-t)\,\theta,$$

and N_p is the same expression in the A's and B's. In particular, for $p = 1, 2$, these expressions can be simplified by b, yielding

$$M_1 - N_1 = [a_{2,t}\,(b_{3,m} - B_{3,m}) - a_{1,t}\,(a_{3,m} - A_{3,m})] \cos (m-t)\,\theta$$
$$+ [- a_{1,t}\,(b_{3,m} - B_{3,m}) - a_{2,t}\,(a_{3,m} - A_{3,m})] \sin (m-t)\,\theta,$$
$$M_2 - N_2 = [- a_{1,t}\,(b_{3,m} - B_{3,m}) - a_{2,t}\,(a_{3,m} - A_{3,m})] \cos (m-t)\,\theta$$
$$+ [-a_{2,t}\,(b_{3,m} - B_{3,m}) + a_{1,t}\,(a_{3,m} - A_{3,m})] \sin (m-t)\,\theta,$$

whence

$$\sum_{j=1}^{2} (X_j - Y_j)^2\, t^2 \sum_{s=1}^{2} a_{s,t}^2 = m^2 r^{2m-2t} [(a_{3,m} - A_{3,m})^2 + (b_{3,m} - B_{3,m})^2 + O(r)].$$

Consequently, by (32), there is a neighborhood of P in which, except at P, $\sum_{j=1}^{2} (X_j - Y_j)^2 \neq 0$. This contradiction of hypothesis establishes the theorem.

THE RICE INSTITUTE.

INEQUALITIES AMONG THE INVARIANTS OF PFAFFIAN SYSTEMS

By Donald C. Dearborn

1. Introduction. Associated with any pfaffian system

$$S\colon \omega^\alpha = a_i^\alpha \, dx^i = 0 \qquad (\alpha = 1, 2, \cdots, r; i = 1, 2, \cdots, n)$$

are certain arithmetic invariants. Among these are the number r of independent equations in the system, the species σ, the class p, and the half-rank ρ.[1] These invariants are all non-negative integers.

The object of this paper is to find sets of inequalities which must be satisfied by these four invariants for any pfaffian system. If for every non-negative integral solution of such a set of inequalities it is possible to find a pfaffian system having that solution as its invariants, the set of inequalities will be called *complete*.

In §2 sets of inequalities are found which hold for any pfaffian system. These sets are not, in general, complete sets. In §3 a complete set of inequalites is given for systems having equal species and half-rank. Included in this classification are all completely separable[2] systems, such as passive systems, systems consisting of a single equation, and systems having $r - 1$ integrals. Systems having rank two are considered in §4. It is shown that such systems have species one or two, and complete sets of inequalities are obtained.

2. Inequalities satisfied by the invariants of any system. It is known that[3] $\rho \leqq \sigma$ and that[4] $p \geqq r + \sigma + 1$ unless the system is passive.

Since there are r independent equations in S, the system may be solved algebraically for r of the differentials and put in the reduced form[5]

$$(2.1) \qquad \omega^\alpha = dx^\alpha + A_\lambda^\alpha \, dx^\lambda \qquad (\alpha = 1, \cdots, r; \lambda = r + 1, \cdots, r + \sigma),$$

where S is assumed to be expressed in terms of the minimum number of differentials. The derived forms are then $\omega'^\alpha = dA_\lambda^\alpha \, dx^\lambda$, which we write as

Received June 4, 1936; part of a doctoral dissertation presented June, 1936 at Duke University.

[1] For definition of class see E. Goursat, *Leçons sur le Problème de Pfaff*, Paris, 1922, p. 268. For species see J. M. Thomas, *Pfaffian systems of species one*, Trans. Amer. Math. Soc., vol. 35 (1933), pp. 356–371. For half-rank see E. Cartan, *Invariants Intégraux*, Paris, 1922, p. 59; Mabel Griffin, *Invariants of pfaffian systems*, Trans. Amer. Math. Soc., vol. 35 (1933), p. 931.

[2] Griffin, loc. cit., p. 936.

[3] J. M. Thomas, *A lower limit for the species of a pfaffian system*, Proc. Nat. Acad. Sci., vol. 19 (1933), p. 913.

[4] Thomas, loc. cit., footnote 1.

[5] Thomas, loc. cit., footnote 1, p. 362.

$(\partial A_\lambda^\alpha + \delta A_\lambda^\alpha)\, dx^\lambda$. Here δA_λ^α represents the differential of A_λ^α formed on the assumption that $x^1, x^2, \cdots, x^{r+\sigma}$ are constant.

The condition that the half-rank is ρ implies that

$$\omega^1 \cdots \omega^r(\partial A_{\lambda_1}^{\alpha_1} + \delta A_{\lambda_1}^{\alpha_1})\, dx^{\lambda_1} \cdots (\partial A_{\lambda_\rho}^{\alpha_\rho} + \delta A_{\lambda_\rho}^{\alpha_\rho})\, dx^{\lambda_\rho} \neq 0$$

for some set of values $\alpha_1, \cdots, \alpha_\rho$. As all such products are of degree $r + 2\rho$, we have the inequality $p \geqq r + 2\rho$.

In order that the half-rank of the system be ρ, it is also necessary that all products

$$\Omega^{\alpha_1 \cdots \alpha_{\rho+1}} \equiv \omega^1 \cdots \omega^r(\partial A_{\lambda_1}^{\alpha_1} + \delta A_{\lambda_1}^{\alpha_1})\, dx^{\lambda_1} \cdots (\partial A_{\lambda_{\rho+1}}^{\alpha_{\rho+1}} + \delta A_{\lambda_{\rho+1}}^{\alpha_{\rho+1}})\, dx^{\lambda_{\rho+1}} \tag{2.2}$$

vanish. Any $\Omega^{\alpha_1 \cdots \alpha_{\rho+1}}$ contains a sum of terms

$$\omega^1 \cdots \omega^r\, dx^{\lambda_1} \cdots dx^{\lambda_{\rho+1}}\, \delta A_{\lambda_1}^{\alpha_1} \cdots \delta A_{\lambda_{\rho+1}}^{\alpha_{\rho+1}}$$

and this sum must vanish since any term from (2.2) involving a ∂A is of degree at least $r + \rho + 2$ in $dx^1, \cdots, dx^{r+\sigma}$. If $\sigma = \rho$, these terms vanish identically. Consequently we suppose that $\sigma \geqq \rho + 1$. In this sum are terms of the type

$$dx^1 \cdots dx^r\, dx^{a_1} \cdots dx^{a_{\rho+1}}\, \delta_{a_1 \cdots a_{\rho+1}}^{l_1 \cdots l_{\rho+1}}\, \delta A_{l_1}^{\alpha_1} \cdots \delta A_{l_{\rho+1}}^{\alpha_{\rho+1}},$$

where $a_1, \cdots, a_{\rho+1}$ is a fixed set taken from $r + 1, \cdots, r + \sigma$ with $a_n \neq a_m$, and $\delta_{a_1 \cdots a_{\rho+1}}^{l_1 \cdots l_{\rho+1}}$ is the generalized Kronecker delta. Since the δA's contain no differentials with index less than $r + \sigma + 1$ and $dx^1 \cdots dx^r\, dx^{a_1} \cdots dx^{a_{\rho+1}} \neq 0$,

$$\delta_{a_1 \cdots a_{\rho+1}}^{l_1 \cdots l_{\rho+1}}\, \delta A_{l_1}^{\alpha_1} \cdots \delta A_{l_{\rho+1}}^{\alpha_{\rho+1}} = 0. \tag{2.3}$$

This last group of terms may be interpreted as the determinant made up of the rows with indices α and the columns with indices a in the matrix $\| \delta A_\lambda^\alpha \|$. Since the elements are subject to the non-cummutative law $\delta A \delta B = -\delta B \delta A$, the same square array gives rise to two different determinants. We shall call the expansion (2.3) the *down-determinant* $| \delta A_\lambda^\alpha |$ and shall say that the matrix $\| \delta A_\lambda^\alpha \|$ has *down-rank* equal to ρ if every down-determinant of order $\rho + 1$ taken from $\| \delta A_\lambda^\alpha \|$ is zero while some down-determinant of order ρ is not zero. Here it is understood that determinants containing any number of repeated rows are to be considered.

A necessary condition that the half-rank of (2.1) be ρ is that the matrix $\| \delta A_\lambda^\alpha \|$ have down-rank $\leqq \rho$. Consider in particular the determinant in which $\alpha_1 = \alpha_2 = \cdots = \alpha_{\rho+1} = \alpha$. The down-determinant equated to zero gives $\delta A_{l_1}^\alpha \delta A_{l_2}^\alpha \cdots \delta A_{l_{\rho+1}}^\alpha = 0$, for all $l_1, l_2, \cdots, l_{\rho+1}$; that is, any δA_λ^α may be expressed as a linear combination of at most ρ of those on the same row. This being true for all values of α, there are at most $r\rho$ independent δA's. Since the system may be expressed in terms of the corresponding A's and $x^1, \cdots, x^{r+\sigma}$, the class p does not exceed $r + r\rho + \sigma$.

A better upper limit for the class may be obtained as follows. Any ω'^α may be written in the form

$$\omega'^\alpha = \eta_\beta^\alpha\, \omega^\beta + \varphi_1^\alpha \varphi_2^\alpha + \cdots + \varphi_{2s_\alpha-1}^\alpha \varphi_{2s_\alpha}^\alpha \quad (s_\alpha \leqq \rho,\ \alpha \text{ not summed}).$$

Since some $\Omega^{\alpha_1 \cdots \alpha_\rho}$ does not vanish, there are 2ρ of the φ's such that

$$\omega^1 \cdots \omega^r \varphi^{\alpha_1}_{2\,l_1-1}\varphi^{\alpha_1}_{2\,l_1} \cdots \varphi^{\alpha_\rho}_{2\,l_\rho-1}\varphi^{\alpha_\rho}_{2\,l_\rho} \neq 0.$$

As every Ω with $\rho + 1$ indices is zero, the products

$$\omega^1 \cdots \omega^r \varphi^{\alpha_1}_{2\,l_1-1}\varphi^{\alpha_1}_{2\,l_1} \cdots \varphi^{\alpha_\rho}_{2\,l_\rho-1}\varphi^{\alpha_\rho}_{2\,l_\rho} \varphi^{\beta}_{2\,l-1}\varphi^{\beta}_{2\,l} \qquad \text{(no sum)}$$

must vanish for every pair β, l; that is, every $\varphi^{\beta}_{2\,l}$ is expressible as a linear combination of the forms

$$\omega^1, \quad \cdots, \quad \omega^r, \quad \varphi^{\alpha_1}_{2\,l_1-1}, \quad \varphi^{\alpha_1}_{2\,l_1}, \quad \cdots, \quad \varphi^{\alpha_\rho}_{2\,l_\rho-1}, \quad \varphi^{\alpha_\rho}_{2\,l_\rho}, \quad \varphi^{\beta}_{2\,l-1}.$$

Hence there are at most $r + r\rho + \rho$ independent forms. The number of equations in the associated set of Ω, Ω^α cannot exceed this number and consequently $p \leqq r + r\rho + \rho$. It is evident that this limit holds in case $\sigma = \rho$.

If we designate by max (a, b) the greater of a and b, we may state

THEOREM 2.1. *The invariants of any pfaffian system satisfy the inequalities*

$$\rho \leqq \sigma, \qquad \max\,(r + 2\rho,\, r + \sigma + 1) \leqq p \leqq r + r\rho + \rho.$$

3. **Systems for which the species and half-rank are equal.** In this section we shall prove

THEOREM 3.1. *If the species and half-rank are equal, then*

$$r + \sigma + \rho \leqq p \leqq r + r\rho + \rho$$

is a complete set of inequalities.

The result is clearly true for passive systems ($\rho = \sigma = 0$).

When the species and half-rank are equal and different from zero, it is clear that $r + \sigma + \rho = r + 2\rho \geqq r + \rho + 1$. Theorem 2.1 then shows that the class p satisfies the inequality of Theorem 3.1. It remains to be shown that for any positive integers r and ρ, there exists at least one pfaffian system of r equations whose class is any integer p satisfying the inequalities of the theorem and whose half-rank and species are the given number ρ.

Any integer p satisfying the inequality of the theorem may be written as $r + \beta\rho + \tau$, $2 \leqq \beta \leqq r$, $0 \leqq \tau \leqq \rho$. Let $x^1, x^2, \cdots, x^{r+\beta\rho+\tau}$ be a set of independent variables, and form the pfaffian system

$$\begin{aligned}
\omega^1 &= dx^1 + x^{r+\rho+1}\,dx^{r+1} + \cdots + x^{r+2\rho}\,dx^{r+\rho} = 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\omega^{\beta-1} &= dx^{\beta-1} + x^{r+(\beta-1)\rho+1}\,dx^{r+1} + \cdots + x^{r+\beta\rho}\,dx^{r+\rho} = 0,\\
\omega^{\beta} &= dx^{\beta} + x^{r+\beta\rho+1}\,dx^{r+1} + \cdots + x^{r+\beta\rho+\tau}\,dx^{r+\tau} = 0,\\
\omega^{\beta+1} &= dx^{\beta+1} = 0,\\
&\cdots\cdots\cdots\cdots\\
\omega^{r} &= dx^{r} = 0.
\end{aligned}$$

This is a system of r independent equations. It has species at most ρ, since it is expressed in terms of $r + \rho$ differentials. To determine the rank we examine the product

$$\begin{aligned}\Omega[\omega'^1]^\rho &= \omega^1 \cdots \omega^r[dx^{r+\rho+1}\, dx^{r+1} + \cdots + dx^{r+2\rho}\, dx^{r+\rho}]^\rho \\ &= \pm\, \underline{|\rho}\; dx^1 \cdots dx^r\, dx^{r+1} \cdots dx^{r+\rho}\, dx^{r+\rho+1} \cdots dx^{r+2\rho}.\end{aligned}$$

Since the x's are independent variables, this product is different from zero and consequently the half-rank is at least ρ. Combining this with the result that the species is at most ρ and the inequality, half-rank $\leqq$ species, of Theorem 2.1, we have that the half-rank and species are each ρ.

The characteristic system consists of the equations $dx^1 = 0, \cdots, dx^{r+\beta\rho+\tau} = 0$, and since these form an independent set of equations, the class is $r + \beta\rho + \tau$. The inequalities of Theorem 3.1 thus form a complete set.

As pointed out above, this case includes all completely separable systems. Accordingly if the species exceeds the half-rank, the number of equations is always greater than one and the number of integrals less than $r - 1$.

4. **Pfaffian systems of rank two.** Next to passive systems, the simplest systems are those for which the rank is two. From the results of §3 we know that if the species is one, $r + 2 \leqq p \leqq 2r + 1$ is a complete set of inequalities. We shall show in this section that if the rank is two, the species cannot exceed two, and that if both rank and species are two, the class is exactly $r + 3$.

We first prove

LEMMA 1. *If $\rho = 1$ and $\sigma \geqq 2$, then $p = r + \sigma + 1$.*

Suppose the system is written in the form (2.1). Then from §2 we know that at least one $\delta A^\alpha_\lambda \neq 0$ and that the matrix $\| \delta A^\alpha_\lambda \|$ has down-rank one. By renumbering the equations and variables δA^1_{r+1} can be made different from zero. If $\delta A^1_{r+1}\, \delta A^\alpha_\lambda = 0$ for all values of α and λ, the class is exactly $r + \sigma + 1$. From §2 we know that $\delta A^\alpha_{\lambda_1} \delta A^\alpha_{\lambda_2} = 0$ (no sum) for all α, λ_1, λ_2; that is, if for some λ_1, $\delta A^\alpha_{\lambda_1} \neq 0$, we have $\delta A^\alpha_{\lambda_2} = \eta^\alpha_{\lambda_2} \delta A^\alpha_{\lambda_1}$ (no sum) for all λ_2. Suppose for some $\alpha \neq 1$ there is a δA^α_λ for which $\delta A^1_{r+1}\, \delta A^\alpha_\lambda \neq 0$. The down-determinant

$$\begin{vmatrix} \delta A^1_{r+1} & \delta A^1_\lambda \\ \delta A^\alpha_{r+1} & \delta A^\alpha_\lambda \end{vmatrix} = \begin{vmatrix} \delta A^1_{r+1} & \eta^1_\alpha\, \delta A^1_{r+1} \\ \delta A^\alpha_{r+1} & \delta A^\alpha_\lambda \end{vmatrix}$$

must vanish. Since $\delta A^1_{r+1}\, \delta A^\alpha_\lambda$ is not zero, $\delta A^1_{r+1}\, \delta A^\alpha_{r+1}$ is not. Consequently all the independent δA's may be taken in the first column.

Now consider any two rows of the matrix $\| \delta A^\alpha_\lambda \|$ for which

$$\delta A^\alpha_{r+1}\, \delta A^\beta_{r+1} \neq 0.$$

These may be written

$$\left\| \begin{matrix} \delta A^\alpha_{r+1} & \eta^\alpha_{r+2}\, \delta A^\alpha_{r+1} & \cdots & \eta^\alpha_{r+\sigma}\, \delta A^\alpha_{r+1} \\ \delta A^\beta_{r+1} & \eta^\beta_{r+2}\, \delta A^\beta_{r+1} & \cdots & \eta^\beta_{r+\sigma}\, \delta A^\beta_{r+1} \end{matrix} \right\| \qquad \text{(no sum).}$$

Since every down-determinant of order two must be zero, we have $\eta_\lambda^\alpha = \eta_\lambda^\beta$ for all λ. This must hold for every pair α, β for which $\delta A_{r+1}^\alpha \, \delta A_{r+1}^\beta \neq 0$. Hence if there is any such pair, all the η's are independent of α, except for those rows on which all the δA's are identically zero.

Suppose now that the system (2.1) has half-rank 1, species $\sigma \geqq 2$, and that there are more than one independent δA's. We shall show that it is possible to adjoin to the system $\sigma - 1$ equations giving a passive system. This will contradict the hypothesis that the species is σ. We adjoin first the equations

$$dx^{r+3} = dx^{r+4} = \cdots = dx^{r+\sigma} = 0.$$

The Ω^α's for this system will be of the form

$$\omega^1 \cdots \omega^r \,(\partial A_\lambda^\alpha \, dx^\lambda dx^{r+3} \cdots dx^{r+\sigma} + \delta A_\lambda^\alpha \, dx^\lambda dx^{r+3} \cdots dx^{r+\sigma}).$$

A single equation may now be added which will cause these to vanish and thus render the system passive. To demonstrate this, we adjoin the equation $dy = 0$, where y is an undetermined function of $x^1, x^2, \cdots, x^{r+\sigma}$ only. The terms containing $\partial A_\lambda^\alpha$ will vanish, since they will be of degree $r + \sigma + 1$ in $r + \sigma$ differentials. The function y must then be determined to satisfy the equations

$$\omega^1 \cdots \omega^r(\delta A_{r+1}^\alpha \, dx^{r+1} \, dx^{r+3} \cdots dx^{r+\sigma} + \delta A_{r+2}^\alpha \, dx^{r+2} \, dx^{r+3} \cdots dx^{r+\sigma}) \, dy = 0.$$

For those values of α for which $\delta A_{r+1}^\alpha = 0$, there will be no equation. In the remaining equations δA_{r+2}^α may be replaced by $\eta_{r+2} \, \delta A_{r+1}^\alpha$ as indicated above. Since the terms δA_{r+1}^α contain none of the differentials $dx^1, \cdots, dx^{r+\sigma}$, they may be factored out of the equations and every equation in the set then reduces to the single one

$$\omega^1 \cdots \omega^r(dx^{r+1} \, dx^{r+3} \cdots dx^{r+\sigma} + \eta_{r+2} \, dx^{r+2} \, dx^{r+3} \cdots dx^{r+\sigma}) \, dy = 0.$$

This gives rise to the partial differential equation

$$\frac{\partial y}{\partial x^{r+2}} - A_{r+2}^\alpha \frac{\partial y}{\partial x^\alpha} - \eta_{r+2} \left(\frac{\partial y}{\partial x^{r+1}} - A_{r+1}^\alpha \frac{\partial y}{\partial x^\alpha} \right) = 0.$$

The desired function y must also satisfy the inequation

$$\omega^1 \cdots \omega^r \, dx^{r+3} \cdots dx^{r+\sigma} \, dy \neq 0.$$

This requires that one of the expressions

$$\frac{\partial y}{\partial x^{r+2}} - A_{r+2}^\alpha \frac{\partial y}{\partial x^\alpha}, \qquad \frac{\partial y}{\partial x^{r+1}} - A_{r+1}^\alpha \frac{\partial y}{\partial x^\alpha}$$

be different from zero. Since $\eta_{r+2} \neq 0$, a solution will exist and that solution will render the system passive. It follows from this contradiction that if $\sigma \geqq 2$, there can be only one independent δA_λ^α and consequently $p = r + \sigma + 1$.

With the aid of Lemma 1 and a result of Miss Griffin, we may now prove

LEMMA 2. *If* $\rho = 1$, *then* $(\sigma - 1)(\sigma - 2) = 0$.

Miss Griffin has shown[6] that any system for which $\rho = 1$ may be put in a form for which either

$$
(4.1)\qquad
\begin{aligned}
\omega'^{\epsilon} &\equiv 0 && \operatorname{mod}\,(\omega^1, \cdots, \omega^r) \quad (\epsilon = 1, 2, \cdots, r - 3),\\
\omega'^{r-2} &\equiv \varphi_1\varphi_2 && \operatorname{mod}\,(\omega^1, \cdots, \omega^r),\\
\omega'^{r-1} &\equiv \varphi_2\varphi_3 && \operatorname{mod}\,(\omega^1, \cdots, \omega^r),\\
\omega'^{r} &\equiv \varphi_3\varphi_1 && \operatorname{mod}\,(\omega^1, \cdots, \omega^r),
\end{aligned}
$$

or

$$
(4.2)\qquad
\begin{aligned}
\omega'^{\epsilon} &\equiv 0 && \operatorname{mod}\,(\omega^1, \cdots, \omega^r) \quad (\epsilon = 1, 2, \cdots, r^1),\\
\omega'^{\lambda} &\equiv \varphi\psi^{\lambda} && \operatorname{mod}\,(\omega^1, \cdots, \omega^r) \quad (\lambda = r^1 + 1, \cdots, r).
\end{aligned}
$$

The first type is of class $r + 3$ and therefore has species $\leqq 2$. The second type will be shown to have species $\leqq 2$.

If φ' vanishes by virtue of the original system augmented by the equation $\varphi = 0$, adjunction of this equation will render the system passive and show that the species is one. Suppose then that $\omega^1 \cdots \omega^r\varphi\varphi' \neq 0$. The congruences (4.2) imply that $\omega'^{\lambda} = \varphi\psi^{\lambda} + \eta^{\lambda}_{\alpha}\, \omega^{\alpha}$. Since the derived form of a derived form is identically zero, we have

$$(4.3)\qquad [\omega'^{\lambda}]' = \varphi'\psi^{\lambda} - \varphi\psi'^{\lambda} + \eta'^{\lambda}_{\alpha}\, \omega^{\alpha} - \eta^{\lambda}_{\alpha}\, \omega'^{\alpha} = 0 \qquad (\text{all } \lambda).$$

Multiplying by $\omega^1 \cdots \omega^r\varphi$ gives

$$(4.4)\qquad \omega^1 \cdots \omega^r\varphi\varphi'\psi^{\lambda} = 0 \qquad (\text{all } \lambda);$$

that is

$$(4.5)\qquad \varphi' \equiv \psi^{\lambda}\theta^{\lambda} \operatorname{mod}\,(\omega^1, \cdots, \omega^r, \varphi) \qquad (\text{no sum}).$$

The product $\psi^{\lambda}\theta^{\lambda}$ is not zero. Since (4.4) and (4.5) hold for all values of λ, $\varphi' \equiv \psi^{\lambda}\psi^{\mu}$. This implies that there can be at most two independent ψ's and consequently the class is at most $r + 3$ and the species at most two. It should be noted that if the species is two there are always at least $r - 2$ Ω^{α}'s which vanish.

Lemmas 1 and 2 now furnish the proof of

THEOREM 4.1. *The invariants of a pfaffian system of rank two satisfy one or the other of the relations*

(a) $\sigma = 1, \qquad r + 2 \leqq p \leqq 2r + 1;$

(b) $\sigma = 2, \qquad p = r + 3, \qquad r > 1;$

and in either case the set of inequalities is complete.

That the invariants satisfy one or the other of these relations is shown by the

[6] Griffin, loc. cit., p. 931.

lemmas. The necessity of the inequality $r > 1$ in case (b) was pointed out at the close of §3. In case (a) the completeness of the set is given by Theorem 3.1. It remains to be shown that for any integer $r > 1$, there exists at least one system having r equations, $\rho = 1$, $\sigma = 2$, and $p = r + 3$. The system

$$\omega^1 = dx^1 + x^5\, dx^3 + x^2\, dx^4 = 0,$$

$$\omega^2 = dx^4 + x^5\, dx^2 + x^1\, dx^3 = 0,$$

$$\omega^i = dx^i = 0 \qquad (i = 6, 7, \cdots, r + 3)$$

satisfies these conditions and thus shows the completeness of the set (b).

A similar theorem for any value of the rank would be of considerable interest, for in most cases it is easier to compute the rank and class than to compute the species. It is likely that the species never exceeds $\rho + 1$ and that when it has this value $p = r + \sigma + \rho$, but no proof of this conjecture has been obtained.

DUKE UNIVERSITY.

ON THE FOURIER TRANSFORMS OF DISTRIBUTIONS ON CONVEX CURVES

By E. K. Haviland and Aurel Wintner

In a previous paper,[1] the asymptotic formula for the Bessel function J_0 has been applied to the derivation of smoothness properties of infinite convolutions of circular equidistributions. For this purpose, not the asymptotic formula but merely an appraisal was needed. It has been indicated in that paper[2] that the same method is valid also in the case of infinite convolutions of certain distribution functions along convex curves—in particular, in the case of some asymptotic distribution problems connected with the Riemann zeta function. The necessary appraisal of the function corresponding in this more general case to the function J_0 has then been carried through[3] by a simple application of a lemma of van der Corput and Landau.[4] The object of the present paper is to replace this appraisal by an asymptotic formula. While the former corresponds to $J_0(r) = O(r^{-\frac{1}{2}})$, $r \to \infty$, the latter will be a generalization of

$$J_0(r) = r^{-\frac{1}{2}}(2/\pi)^{\frac{1}{2}} \cos (r - \pi/4) + O(r^{-1}), \qquad r \to \infty.$$

The general function to be considered is, in contrast to the particular function $J_0(r)$, a function not only of r but of an angular parameter ψ also. For a fixed value of ψ, the asymptotic formula in question may be obtained by applying an elementary method.[5] What is needed for the applications mentioned above, and what will be proved in what follows, is the fact that the asymptotic formula holds uniformly for all values of ψ, i.e., that the error term is in absolute value less than Cr^{-1}, where C is a constant independent both of r and of ψ.

Let $x = x(\varphi)$, $y = y(\varphi)$, where $0 \leqq \varphi < 2\pi$, be a parametric representation of a convex Jordan curve S in the (x, y)-plane. It will be described more precisely below. Let $\sigma = \sigma(E)$ be an absolutely additive set function defined, for every Borel set E of the (x, y)-plane, by setting $\sigma(F)$ equal to $1/(2\pi)$ times the linear measure of those φ for which $(x(\varphi), y(\varphi))$ is contained in FS, if F is any open set in plane. In particular, it is seen that S is the spectrum[6] of σ. By

Received April 2, 1936.

[1] A. Wintner, loc. cit., I. The references are collected at the end of the paper.

[2] A. Wintner, ibid., pp. 328–329.

[3] B. Jessen and A. Wintner, loc. cit., Theorem 12, p. 63.

[4] Cf., e.g., R. Kershner, loc. cit., where further references are given.

[5] Cf. A. Wintner, loc. cit., III, pp. 57–60, where references to the literature are given.

[6] For the definition of the spectrum, together with some properties of spectra, cf. A. Wintner, loc. cit., II, pp. 9–10, and E. K. Haviland, loc. cit., II, pp. 653–654. It has been pointed out by Professor Khintchine that, contrary to statements in these papers, the vectorial sum of two closed sets is necessarily closed only when at least one of these sets is

virtue of the definition of Lebesgue and of Radon integrals, it may be shown[7] that

$$\Lambda(u, v; \sigma) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp[i(ux + vy)]\, d_{xy}\sigma(E) \tag{1}$$
$$= \frac{1}{2\pi} \int_0^{2\pi} \exp[i(ux(\varphi) + vy(\varphi))]\, d\varphi.$$

On setting $u = r \cos \psi$ and $v = r \sin \psi$, one obtains

$$\Lambda = \Lambda(r \cos \psi, r \sin \psi; \sigma) = \frac{1}{2\pi} \int_0^{2\pi} \exp[irh(\varphi; \psi)]\, d\varphi, \tag{2}$$

where

$$h(\varphi; \psi) = x(\varphi) \cos \psi + y(\varphi) \sin \psi. \tag{2a}$$

It will be assumed that

(i) $x(\varphi)$ and $y(\varphi)$ possess continuous fourth derivatives;

(ii) $h''(\varphi; \psi)$ has for any fixed ψ exactly two zeros on the curve S and these zeros are both simple.[8] Here and in what follows a prime denotes partial differentiation with respect to φ.

Under these assumptions, it will be shown that

$$\Lambda = (2\pi r)^{-\frac{1}{2}}\{[h''(\varphi_3(\psi); \psi)]^{-\frac{1}{2}} \exp\,[i(rh(\varphi_3(\psi); \psi) + \pi/4)] \tag{3}$$
$$+ [-h''(\varphi_1(\psi); \psi)]^{-\frac{1}{2}} \exp\,[i(rh(\varphi_1\psi); \psi) - \pi/4)]\} + O(r^{-1}),$$

where the O-term holds uniformly for all ψ, and $\varphi_1 = \varphi_1(\psi)$ and $\varphi_3 = \varphi_3(\psi)$ denote the two zeros of h' on S. That, for any fixed ψ, there are precisely two such zeros and that they separate the zeros of h'' is a consequence[3] of (ii).

The proof of (3) proceeds as follows. First, the minimum distance between a zero of h' and a zero of h'' has, for reasons of continuity, a positive lower bound ξ independent of ψ. If ψ is fixed, one of the zeros φ of $h'(\varphi; \psi)$, say $\varphi = \varphi_1 = \varphi_1(\psi)$, corresponds to a maximum of the function $h(\varphi; \psi)$, the other, say $\varphi = \varphi_3 = \varphi_3(\psi)$, to a minimum. Let $\varphi_2(\psi)$ and $\varphi_4(\psi)$ be the zeros of $h''(\varphi; \psi)$ and let them be so situated that $\varphi_1 < \varphi_2 < \varphi_3 < \varphi_4 < \varphi_1 + 2\pi$. Finally, let four numbers η_i be so chosen that

$$\varphi_1 < \eta_1 < \varphi_2 < \eta_2 < \varphi_3 < \eta_3 < \varphi_4 < \eta_4 < \varphi_1 + 2\pi.$$

---

bounded. Correspondingly, what is actually proved, loc. cit., is not that the spectrum of the convolution of two distribution functions is the vectorial sum of the spectra of the individual distribution functions, but that it is the closure of this vectorial sum. This does not at all affect the validity of the proofs given in the papers referred to.

[7] Cf. E. K. Haviland, loc. cit., I, pp. 552–553. The reasoning there used in the proof of Theorem V may be used unchanged to prove equation (1) of the present paper.

[8] (ii) might be generalized under suitable assumptions to the case where the second derivative has more than two zeros and has multiple zeros.[5] The case treated here is the one of interest for applications to infinite convolutions of the type occurring in the distribution theory of the Riemann zeta function.

As $h'''(\varphi; \psi)$ is continuous on the torus T: $(0 \leqq \varphi < 2\pi; 0 \leqq \psi < 2\pi)$, $| h'''(\varphi; \psi) |$ is bounded, say $< M$, there. Moreover, it is clear from (ii) that there exists a positive constant α such that $| h''(\varphi_1(\psi); \psi) | > \alpha$ for every ψ. Then one may choose $\eta_1(\psi)$ so that it lies between $\varphi_1(\psi) + \frac{1}{3}\zeta$ and $\varphi_1(\psi) + \frac{2}{3}\zeta$ for all ψ, where $\zeta = \min(\xi, \alpha/(2M))$ is independent of ψ, and a similar choice will be made for the other three η_i's.

From (2)

$$\Lambda = \frac{1}{2\pi}\left\{\int_{\varphi_1}^{\eta_1} + \int_{\eta_1}^{\eta_2} + \int_{\eta_2}^{\varphi_2} + \int_{\varphi_2}^{\eta_3} + \int_{\eta_3}^{\eta_4} + \int_{\eta_4}^{\varphi_1+2\pi}\right\} \tag{4}$$
$$\equiv J_I + J_{II} + J_{III} + J_{IV} + J_V + J_{VI},$$

say.

In order to treat J_I, set, for $\varphi_1 \leqq \varphi \leqq \eta_1$, where $\varphi_1 = \varphi_1(\psi)$ and $\eta_1 = \eta_1(\psi)$,

$$t^2 = h(\varphi_1; \psi) - h(\varphi; \psi) \tag{5}$$

for every fixed ψ, corresponding to the fact that φ_1 is a *simple* zero of h' by assumption (ii). On taking the positive square root,

$$t = | h(\varphi_1; \psi) - h(\varphi, \psi) |^{\frac{1}{2}}, \tag{6}$$

so that as φ increases steadily from φ_1 to η_1, the variable t increases steadily from zero to a quantity $a_1(\psi) = | h(\varphi_1(\psi); \psi) - h(\eta_1(\psi); \psi) |^{\frac{1}{2}}$ which has a positive lower bound β independent of ψ in virtue of assumption (ii).

Moreover, if a dot represents partial differentiation with respect to t,

$$\dot{\varphi} = -2t/h'(\varphi(t, \psi); \psi), \qquad \text{if } 0 < t \leqq a_1(\psi), \tag{7}$$

so

$$J_I = -\frac{1}{\pi}\exp[irh(\varphi_1(\psi); \psi)]\int_0^{a_1(\psi)} \exp[-irt^2]\, t/h'(\varphi(t, \psi); \psi)\, dt. \tag{8}$$

The integral in (8) is of the form $\int_0^{a_1(\psi)} f(t; \psi)\exp[-irt^2]\,dt$ and may accordingly, for every *fixed* ψ, be evaluated asymptotically in a known manner[5] under the assumption that $f(t; \psi)$ possesses a continuous second partial derivative with respect to t in $0 \leqq t \leqq a_1(\psi)$. That the function

$$f(t; \psi) = t/h'(\varphi(t, \psi); \psi) \tag{9}$$

possesses this last property may be seen as follows.

By Taylor's Theorem with the integral form of the remainder

$$h(\varphi; \psi) = h(\varphi) = h(\varphi_1) + (\varphi - \varphi_1)h'(\varphi_1) + \tfrac{1}{2}(\varphi - \varphi_1)^2 h''(\varphi_1)$$
$$+ \tfrac{1}{2}\int_0^{\varphi-\varphi_1} h'''(\varphi_1 + s)(\varphi - \varphi_1 - s)^2\, ds, \tag{10}$$

where $\varphi_1 = \varphi_1(\psi)$. Since $h'(\varphi) = 0$, we have from (6), on changing the integration variable in (10),

$$t = \left\{-\tfrac{1}{2}(\varphi - \varphi_1)^2 h''(\varphi_1) - \tfrac{1}{2}\int_{\varphi_1}^{\varphi} h'''(s)(\varphi - s)^2\,ds\right\}^{\frac{1}{2}}. \tag{11}$$

Applying the same form of Taylor's Theorem to $h'(\varphi, \psi)$, we have

$$h'(\varphi; \psi) = h'(\varphi) = (\varphi - \varphi_1)h''(\varphi_1) + \int_{\varphi_1}^{\varphi} h'''(s)(\varphi - s)\,ds. \tag{12}$$

Now φ is a continuous function of t for fixed ψ, as may be seen from (5), since t is a monotone continuous function of φ for $\varphi_1(\psi) < \varphi \leqq \eta_1(\psi)$ and $h'(\varphi; \psi)$ is negative there and has a *simple* zero at $\varphi = \varphi_1(\psi)$. Then substituting (11) and (12) into (9) and writing the parameter explicitly, we obtain

$$\begin{aligned} f(t) = f(t; \psi) &= \frac{\left\{-\frac{1}{2}h''(\varphi_1(\psi); \psi) - \frac{1}{2}\displaystyle\int_{\varphi_1(\psi)}^{\varphi(t,\psi)} h'''(s; \psi)\left[1 - \frac{s - \varphi_1(\psi)}{\varphi(t, \psi) - \varphi_1(\psi)}\right]^2 ds\right\}^{\frac{1}{2}}}{h''(\varphi_1(\psi); \psi) + \displaystyle\int_{\varphi_1(\psi)}^{\varphi(t,\psi)} h'''(s; \psi)\left[1 - \frac{s - \varphi_1(\psi)}{\varphi(t, \psi) - \varphi_1(\psi)}\right] ds} \\ &= \{-\tfrac{1}{2}(h_1'' + H_{12})\}^{\frac{1}{2}}/\{h_1'' + H_{11}\}, \end{aligned} \tag{13}$$

where

$$H_{ij} = [\varphi(t, \psi) - \varphi_i(\psi)]^{-j}\int_{\varphi_i(\psi)}^{\varphi(t,\psi)} h'''(s; \psi)[\varphi(t, \psi) - s]^j\,ds \quad (i = 1, 3;\ j = 1, 2),$$

and

$$h_i^{(l)} = h^{(l)}(\varphi_i(\psi); \psi) \qquad (i = 1, 3;\ l = 0, 1, 2, 3, 4).$$

Now the H_{ij}, $j = 1, 2$, are, for fixed ψ, continuous functions of φ for $\varphi_1(\psi) < \varphi \leqq \eta_1(\psi)$. Moreover,

$$|H_{ij}| \leqq (\varphi - \varphi_1)\,M, \tag{14}$$

where M is the maximum of $|h'''(\varphi; \psi)|$ on the torus T, so that the H_{ij} are, for fixed ψ, continuous functions of φ in $\varphi_1(\psi) \leqq \varphi \leqq \eta_1(\psi)$ also. Since $0 \leqq \varphi - \varphi_1 < \zeta \leqq \alpha/(2M)$ for all φ in $\varphi_1(\psi) \leqq \varphi \leqq \eta_1(\psi)$, i.e., for all t in $0 \leqq t \leqq a_1(\psi)$, and for all ψ, it follows from (14) that $|H_{ij}| < \frac{1}{2}\alpha$, so that $|h_1'' + H_{11}| > \frac{1}{2}\alpha$ there. Because φ is, for fixed ψ, a continuous function of t in $0 \leqq t \leqq a_1(\psi)$, it is seen that $f(t; \psi)$ possesses this same property. Then as $\dot\varphi = -2f$, it is clear that $\dot\varphi$ tends to a definite limit as $t \to +0$, which implies that $\dot\varphi$ exists, and (7) holds, at $t = 0$ also. Differentiation of (13) with respect to t, substitution for $\dot\varphi$ from (7) and (13), and suitable grouping of the factors $(\varphi - \varphi_1)^{-1}$ in the result gives

$$\dot f = \{h_1'' + H_{11}\}^{-3}\{[h_1'' + H_{11}][L_{11} - L_{12}] - [h_1'' + H_{12}]L_{11}\}, \tag{15}$$

where

$$L_{ik} = [\varphi(t, \psi) - \varphi_i(\psi)]^{-(k+1)}\int_{\varphi_i(\psi)}^{\varphi(t,\psi)} h'''(s; \psi)[s - \varphi_i(\psi)]^k\,ds \quad (i = 1, 3;\ k = 1, 2, 3).$$

In a similar manner, one obtains

$$
(16)\quad
\begin{aligned}
\dddot{f} = \{h_1'' + H_{11}\}^{-5}\{-\tfrac{1}{2}(h_1'' + H_{12})\}^{\frac{1}{2}}\{[h_1'' + H_{11}][-2L_{11}(L_{10} - L_{11})\\
+ 3L_{12}(L_{10} - L_{11}) + h_1''Q - (L_{10} - 2L_{11} + L_{12})h'''\\
+ 2h_1''L_{11}(L_{10} - 2L_{11} + L_{12}) - L_{11}(L_{11} - L_{12})]\\
- 3L_{11}[H_{11}L_{11} - H_{11}L_{12} - h_1''L_{12} - H_{12}L_{11}]\},
\end{aligned}
$$

where

$$
Q = \frac{3}{[\varphi(t, \psi) - \varphi_1(\psi)]^4}\int_{\varphi_1(\psi)}^{\varphi(t,\psi)} h'''(s;\psi)[s - \varphi_1(\psi)]^2 ds - \frac{h'''(\varphi(t, \psi); \psi)}{\varphi(t, \psi) - \varphi_1(\psi)}.
$$

Now the L_{1k}, $k = 1, 2, 3$, are, for fixed ψ, continuous functions of φ for $\varphi_1(\psi) < \varphi \leqq \eta_1(\psi)$. In addition, it follows from the continuity of $h'''(\varphi)$ at $\varphi = \varphi_1 = \varphi_1(\psi)$ that

$$
h'''(\varphi) = h'''(\varphi_1) + \omega(\varphi); \qquad |\,\omega(\varphi)\,| < \epsilon, \text{ if } |\,\varphi - \varphi_1\,| < \delta.
$$

Then

$$
L_{1k} = \frac{h'''(\varphi_1)}{k + 1} + \frac{1}{(\varphi - \varphi_1)^{k+1}}\int_{\varphi_1}^{\varphi} \omega(s)(s - \varphi_1)^k ds,
$$

so

$$
|\,L_{1k} - h'''(\varphi_1)/(k + 1)\,| < \epsilon, \qquad \text{if } |\,\varphi - \varphi_1\,| < \delta.
$$

Consequently, the L_{1k} are, for fixed ψ, continuous functions of φ in the closed interval $\varphi_1(\psi) \leqq \varphi \leqq \eta_1(\psi)$.

Furthermore,

$$
Q = 3(\varphi - \varphi_1)^{-4}\int_{\varphi_1}^{\varphi} h'''(s)(s - \varphi_1)^2 ds - h'''(\varphi)/(\varphi - \varphi_1).
$$

By partial integration, this becomes

$$
-(\varphi - \varphi_1)^{-4}\int_{\varphi_1}^{\varphi} h^{(iv)}(s)(s - \varphi_1)^3 ds = -(\varphi - \varphi_1)^{-4}\int_{\varphi_1}^{\varphi} [h^{(iv)}(\varphi_1) + \omega_1(s)](s - \varphi_1)^3 ds,
$$

where $|\,\omega_1(\varphi)\,| < \epsilon$ provided $|\,\varphi - \varphi_1\,| < \delta_1$. Then

$$
Q = -\tfrac{1}{4}h^{(iv)}(\varphi_1) - (\varphi - \varphi_1)^{-4}\int_{\varphi_1}^{\varphi} \omega_1(s)(s - \varphi_1)^3 ds,
$$

and the second term of this expression is in absolute value $<\epsilon$, provided $|\,\varphi - \varphi_1\,| < \delta_1$. Hence, for fixed ψ, Q is a continuous function of φ in $\varphi_1(\psi) \leqq \varphi \leqq \eta_1(\psi)$. In addition, the denominators in (15) and (16) are respectively greater in absolute value than $\alpha^3/8$, $\alpha^5/32$, both of which are positive, and these inequalities hold uniformly for all ψ.

Inasmuch as φ is, for fixed ψ, a continuous function of t in $0 \leqq t \leqq a_1(\psi)$, it follows from the preceding remarks and from (i) that

(16a): f, $\ddot{f}$ and $\dddot{f}$ are for fixed ψ, continuous functions of t in $0 \leqq t \leqq a_1(\psi)$.

Finally, $|H_{1j}| < 2\pi M$, $j = 1, 2$; $|L_{1k}| \leqq M$; $|Q| \leqq N$, where N is the maximum value of $|h^{(\text{iv})}(\varphi; \psi)|$ on T; and $|h_1'' + H_{11}| > \frac{1}{2}\alpha$, for all t in $0 \leqq t \leqq a_1(\psi)$, where M, N and α are independent of ψ. Consequently,
(16b): if $0 \leqq t \leqq a_1(\psi)$, f, $\dot f$ and $\ddot f$ are bounded uniformly with respect to ψ.

On placing

(17) $tg(t; \psi) = f(t; \psi) - f(0; \psi) - t\dot f(0; \psi)$, if $0 < t \leqq a_1(\psi)$, and $g(0; \psi) = 0$,

it follows that $g(t; \psi)$ possesses a continuous first partial derivative with respect to t not only for $0 < t \leqq a_1(\psi)$, where this can be seen at once, but for $0 \leqq t \leqq a_1(\psi)$ as well. In particular,

$$f(t; \psi) = f(0; \psi) + \int_0^t \dot f(s; \psi)\, ds,$$

so

$$g(t; \psi) = \frac{1}{t} \int_0^t [\dot f(s; \psi) - \dot f(0; \psi)]\, ds \tag{18}$$

and

$$\begin{aligned} \dot g(t; \psi) &= \frac{1}{t^2} \int_0^t [\dot f(t; \psi) - \dot f(s, \psi)]\, ds \\ &= \frac{1}{t^2} \int_0^t \ddot f(s + \vartheta(t - s); \psi)\, (t-s)\, ds. \end{aligned} \tag{19}$$

The quantity ϑ depends on t, s and ψ, but is such that $0 < \vartheta < 1$ for all $0 \leqq s \leqq t$; $0 \leqq t \leqq a_1(\psi)$; $0 \leqq \psi < 2\pi$. In view of (16a), (19) implies the above-mentioned continuity of $\dot g$ as a function of t, for fixed ψ. Furthermore, (16b), (18) and (19) imply
(20a): $\dot g(t; \psi)$ is for $0 \leqq t \leqq a_1(\psi)$ bounded uniformly with respect to ψ;
(20b): $g(a_1(\psi); \psi)$ is a bounded function of ψ.

Moreover, by (6),

$$a_1(\psi) \leqq \sqrt{2\mu} \tag{21}$$

where μ is the maximum of $|h(\varphi; \psi)|$ on the torus T.

It will now be shown that for every ψ and for every $r > 0$

$$\left| \int_0^{a_1(\psi)} tg(t; \psi) \exp[-irt^2]\, dt \right| < C_1/r, \tag{22}$$

where C_1, like the quantities $C_2, \cdots, C_4$; $C_I, \cdots, C_{VI}$, to be used in what follows, is a constant independent both of ψ and of r. First, on applying partial integration, the integral in (22) may be written in the form

$$g(a_1(\psi); \psi) \int_0^{a_1(\psi)} t \exp[-irt^2]\, dt - \int_0^{a_1(\psi)} \dot g(t; \psi) \int_0^t y \exp[-iry^2]\, dy\, dt$$

or

$$r^{-1}\left[g(a_1(\psi);\psi)\,F(a_1r^{\frac{1}{2}}) - \int_0^{a_1(\psi)} \dot{g}(t;\psi)\,F(tr^{\frac{1}{2}})\,dt\right], \tag{23}$$

where

$$F(t) = \int_0^t y \exp[-iy^2]\,dy, \text{ hence } F(t^2) = \tfrac{1}{2}\int_0^{t^4} e^{-iy}\,dy,$$

from which it is seen that

$$| F | \leqq 1. \tag{24}$$

From (20a), (20b), (21) and (24) it follows that the quantity in brackets in (23) is in absolute value less than a constant C_1. This proves (22).

Again,

$$\left| \int_0^{a_1(\psi)} t \exp[-irt^2]\,dt \right| < C_2/r, \tag{25}$$

for on placing $r^{\frac{1}{2}} = y$, this integral becomes $F(r^{\frac{1}{2}}\,a_1(\psi))/r$, which implies (25) in view of (24).

Finally,

$$\left| \int_{a_1(\psi)}^{+\infty} \exp[-irt^2]\,dt \right| < C_3/r. \tag{26}$$

For $G_2(r) = \int_r^{+\infty} y^{-\frac{1}{2}} \exp[-iy]\,dy$ exists and is $O(r^{-\frac{1}{2}})$ in virtue of the second mean value theorem applied to a finite interval. Setting $rt^2 = y$, the integral in (26) becomes, up to a constant factor,

$$r^{-\frac{1}{2}}G_2\,(r[a_1(\psi)]^2) = O(r^{-1}), \text{ since } a_1(\psi) > \beta > 0 \text{ for all } \psi,$$

where β is the constant defined above, following equation (6).

Substituting (17) and (25) into (22) and combining the result with (26), one obtains, in virtue of (20a) and the fact that $f(0, \psi)$ and $\dot{f}(0, \psi)$ are bounded functions of ψ,

$$\left| \int_0^{a_1(\psi)} f(t;\psi) \exp[-irt^2]\,dt - f(0;\psi) \int_0^{+\infty} \exp[-irt^2]\,dt \right| < C_4/r. \tag{27}$$

Since[9]

$$\int_0^{+\infty} \exp[-ix^2]\,dx = \tfrac{1}{2}\,\pi^{\frac{1}{2}} \exp[-i\pi/4],$$

one obtains from (8), (13) and (27)

$$| J_I - \tfrac{1}{2}\,[-2\pi h''(\varphi_1(\psi);\psi)r]^{-\frac{1}{2}} \exp[i(rh(\varphi_1(\psi);\psi) - \pi/4)] \,| < C_I/r. \tag{28}$$

[9] Cf. e.g., W. F. Osgood, op. cit., pp. 308–309.

To calculate the integral J_{II}, we observe that $h'(\varphi; \psi)$is negative for

$$\eta_1(\psi) \leqq \varphi \leqq \eta_2(\psi),$$

so that $h(\eta_1(\psi); \psi) - h(\varphi; \psi)$ is in this interval steadily increasing from zero, and if we set

$$t = | h(\eta_1(\psi); \psi) - h(\varphi; \psi) |^{\frac{1}{2}},$$

t increases from 0 to $a_2(\psi) = | h(\eta_1(\psi); \psi) - h(\eta_2(\psi); \psi) |^{\frac{1}{2}}$ as φ increases from $\eta_1(\psi)$ to $\eta_2(\psi)$. By the introduction of t as integration variable in J_{II},

$$J_{II} = -\frac{1}{\pi} \exp\,[irh(\eta_1(\psi); \psi)] \int_0^{a_2(\psi)} \exp\,[-irt^2]\, t/h'(\varphi(t, \psi); \psi)\, dt.$$

This last integral is of the form $\int_0^{a_2(\psi)} f(t; \psi) \exp\,[-irt^2]\, dt$, where

$$f = f(t; \psi) = t/h'(\varphi(t, \psi); \psi).$$

Differentiation with respect to t and substitution for $\dot\varphi$ from (7) gives

$$\dot f(t; \psi) = \{h'(\varphi(t, \psi); \psi)\}^{-3}\, \{[h'(\varphi(t, \psi); \psi)]^2 + 2t^2 h''(\varphi(t, \psi); \psi)\}.$$

Thereupon a similar straightforward calculation shows

$$\begin{aligned} \ddot f(t; \psi) = \{h'(\varphi(t, \psi); \psi)\}^{-5}\, \{&-4t^3 h'(\varphi(t, \psi); \psi) h'''(\varphi(t, \psi); \psi) \\ &+ 6t[h'(\varphi(t, \psi); \psi)]^2\, h''(\varphi(t, \psi); \psi) + 12t^3[h''(\varphi(t, \psi); \psi)]^2\}. \end{aligned} \tag{29}$$

Just as $\eta_1(\psi)$ has already been so chosen that $\frac{1}{3}\zeta < \eta_1(\psi) - \varphi_1(\psi) < \frac{2}{3}\zeta$, one may so choose $\eta_2(\psi)$ that $\frac{1}{3}\zeta < \varphi_3(\psi) - \eta_2(\psi) < \frac{2}{3}\zeta$, where ζ is independent of ψ. Then from (i), (29), (2a) and the continuity of φ as a function of t, it is seen that $\ddot f(t; \psi)$ is, for every fixed ψ, a continuous function of t in

$$0 \leqq t \leqq a_2(\psi) \leqq \sqrt{2\mu},$$

where μ, as before, is the maximum of $| h(\varphi; \psi) |$ on the torus T. Moreover $\dot f(t; \psi)$ and $\ddot f(t; \psi)$ are for $0 \leqq t \leqq a_2(\psi)$ uniformly bounded with respect to ψ. If $g(t; \psi)$ be again defined by (17), these facts imply that
(30a): $g(t; \psi)$ is for $0 \leqq t \leqq a_2(\psi)$ uniformly bounded with respect to ψ;
(30b): $g(a_2(\psi); \psi)$ is a bounded function of ψ.

Finally, $f(0; \psi) = 0/h'(\eta_1(\psi); \psi) = 0$ for all ψ. By the same reasoning as that used in the calculation of J_I, it then follows that

$$| J_{II} | < C_{II}/r. \tag{31}$$

To calculate the integral J_{III}, we set

$$t = | h(\varphi; \psi) - h(\varphi_3(\psi); \psi) |^{\frac{1}{2}}.$$

Then

$$J_{III} = \frac{1}{2\pi} \int_{\eta_3(\psi)}^{\varphi_3(\psi)} \exp\,[irh(\varphi; \psi)]\, d\varphi$$

$$= -\frac{1}{\pi} \exp\,[irh(\varphi_3(\psi); \psi)] \int_0^{a_3(\psi)} f(t; \psi) \exp\,[irt^2]\, dt,$$

where

$$f(t; \psi) = t/h'(\varphi(t, \psi); \psi) = -\{\tfrac{1}{2}\,[h_3'' + H_{32}]\}^{1/2}/(h_3'' + H_{31}).$$

A calculation precisely similar to that of J_I then shows that

$$(32) \quad |\, J_{III} - \tfrac{1}{2}\,[2\pi h''(\varphi_3(\psi); \psi) r]^{-\frac{1}{2}} \exp\,[i(rh(\varphi_3(\psi); \psi) + \pi/4)]\,| < C_{III}/r.$$

Similarly, if in J_{IV} we set

$$t = |\, h(\varphi; \psi) - h(\varphi_3(\psi); \psi)\,|^{\frac{1}{2}},$$

we find

$$J_{IV} = \frac{1}{\pi} \exp\,[irh(\varphi_3(\psi); \psi)] \int_0^{a_4(\psi)} f(t; \psi) \exp\,[irt^2]\, dt,$$

where

$$f(t; \psi) = t/h'(\varphi(t, \psi); \psi) = \{\tfrac{1}{2}\,[h_3'' + H_{32}]\}^{\frac{1}{2}}/(h_3'' + H_{31}).$$

A calculation precisely similar to that of I and III then shows that

$$(33) \quad |\, J_{IV} - \tfrac{1}{2}\,[2\pi h''(\varphi_3(\psi); \psi) r]^{-\frac{1}{2}} \exp\,[i(rh(\varphi_3(\psi); \psi) + \pi/4)]\,| < C_{IV}/r.$$

The case of J_V is similar to that of J_{II}. Here we set

$$t = |\, h(\varphi; \psi) - h(\eta_3(\psi); \psi)\,|^{\frac{1}{2}},$$

so that

$$J_V = \frac{1}{\pi} \exp\,[irh(\eta_3(\psi); \psi)] \int^{a_5(\psi)} f(t; \psi) \exp\,[irt^2]\, dt,$$

where $f(t; \psi) = t/h'(\varphi(t, \psi); \psi)$. Since $f(0; \psi) = 0$, it follows, as in the case of J_{II}, that

$$(34) \qquad |\, J_V\,| < C_V/r.$$

Finally, if one sets in J_{VI}

$$t = |\, h(\varphi_1(\psi); \psi) - h(\varphi; \psi)\,|^{\frac{1}{2}},$$

one obtains

$$J_{VI} = \frac{1}{\pi} \exp\,[irh(\varphi_1(\psi); \psi)] \int_0^{a_6(\psi)} f(t; \psi) \exp\,[-irt^2]\, dt,$$

where

$$f(t;\psi) = t/h'(\varphi(t,\psi);\psi) = -\{-\tfrac{1}{2}[h_1'' + H_{52}]\}^{\frac{1}{2}}/[h_1'' + H_{51}], \ (\varphi_5 = \varphi_1 + 2\pi).$$

Then reasoning very similar to that used in the case of J_I shows that

$$(35) \quad | J_{VI} - \tfrac{1}{2}[-2\pi h''(\varphi_1(\psi);\psi)r]^{-\frac{1}{2}} \exp [i(rh(\varphi_1(\psi);\psi) - \pi/4)] | < C_{VI}/r.$$

Combining (28), (31), (32), (33), (34) and (35), we obtain (3) by virtue of (4).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES

E. K. HAVILAND, I, *On statistical methods in the theory of almost-periodic functions*, Proceedings of the National Academy of Sciences, vol. 19 (1933), pp. 549–555.

—— II, *On the theory of absolutely additive distribution functions*, American Journal of Mathematics, vol. 56 (1934), pp. 625–658.

B. JESSEN AND A. WINTNER, *Distribution functions and the Riemann zeta-function*, Transactions of the American Mathematical Society, vol. 38 (1935), pp. 48–88.

R. KERSHNER, *Determination of a van der Corput-Landau absolute constant*, American Journal of Mathematics, vol. 57 (1935), pp. 840–846.

W. F. OSGOOD, *Lehrbuch der Funktionentheorie*, 5th ed., Leipzig, 1928.

A. WINTNER, I, *Upon a statistical method in the theory of diophantine approximations*, American Journal of Mathematics, vol. 55 (1933), pp. 309–331.

—— II, *On the addition of independent distributions*, American Journal of Mathematics, vol. 56 (1934), pp. 8–16.

—— III, *On the asymptotic formulae of Riemann and of Laplace*, Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 57–62.

FUNCTIONS REPRESENTABLE BY TWO LAPLACE INTEGRALS

By D. H. Ballou

1. **Introduction.** One of the properties of the Laplace integral representation of a function $f(z)$,

$$f(z) = \int_0^\infty e^{-zt}\varphi(t)\,dt, \tag{1}$$

is the uniqueness of the determining function $\varphi(t)$ when that function is continuous.[1] It has been pointed out by G. Doetsch[2] that if a function $f(z)$ may be expanded into two different series the terms of which are representable by Laplace integrals and if the term by term transformation of those series is permissible, then this function is represented by two different Laplace integrals. Furthermore, there will follow from the uniqueness property the equality of two new series, the determining functions of the integrands.

It has been known that the cotangent was one function capable of such a representation,[3] for it has series developments both in terms of partial fractions and of exponentials:

$$\operatorname{ctn} z = \frac{1}{z} + 2\sum_{n=1}^{\infty} \frac{z}{z^2 - n^2\pi^2}, \tag{2}$$

$$\operatorname{ctn} z = -i\left(1 + 2\sum_{n=1}^{\infty} e^{2niz}\right). \tag{3}$$

Now if we take the function[4] $\dfrac{\operatorname{ctn}\sqrt{-s}}{-\sqrt{-s}}$ the terms of the series are representable

Received May 11, 1936; presented to the American Mathematical Society, September 1, 1936.

[1] See, for instance, D. V. Widder, *A generalization of Dirichlet's series and Laplace's integrals by means of a Stieltjes integral*, Transactions of the American Mathematical Society, vol. 31 (1929), p. 705.

[2] G. Doetsch, *Überblick über Gegenstand und Methode der Funktionanalysis*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 36 (1927), p. 28.

[3] G. Doetsch, loc. cit. See also H. Hamburger, *Über einige Beziehungen, die mit der Funktionalgleichung der Riemannschen ζ-Funktion äquivalent sind*, Mathematische Annalen, vol. 85 (1922), p. 129.

[4] Here and throughout this paper that branch of the double-valued function $z = \sqrt{-s}$ is taken which corresponds to the upper half of the z-plane. We then are dealing only with single-valued functions and we are assured of the convergence of (5) for $R(s) > 0$.

by Laplace integrals, and the term by term transformation gives us the result that

$$\frac{\operatorname{ctn}\sqrt{-s}}{-\sqrt{-s}} = \frac{1}{s} + 2\sum_{n=1}^{\infty}\frac{1}{s+n^2\pi^2} = \int_0^{\infty} e^{-st}\left[1 + 2\sum_{n=1}^{\infty} e^{-n^2\pi^2 t}\right]dt, \tag{4}$$

$$\frac{\operatorname{ctn}\sqrt{-s}}{-\sqrt{-s}} = \frac{1}{\sqrt{s}} + 2\sum_{n=1}^{\infty}\frac{e^{-2n\sqrt{s}}}{\sqrt{s}} = \int_0^{\infty} e^{-st}\left[\frac{1}{\sqrt{\pi t}} + 2\sum_{n=1}^{\infty}\frac{e^{-\frac{n^2}{t}}}{\sqrt{\pi t}}\right]dt, \tag{5}$$

where the integrals on the right converge for $R(s) > 0$. Thus this form of the cotangent is representable by two Laplace integrals. The uniqueness property establishes the equality of the series of the integrands

$$1 + 2\sum_{n=1}^{\infty} e^{-n^2\pi^2 t} = \frac{1}{\sqrt{\pi t}} + 2\sum_{n=1}^{\infty}\frac{e^{-\frac{n^2}{t}}}{\sqrt{\pi t}}, \tag{6}$$

and this equality proves to be the linear transformation of one of the theta null-functions, ϑ_3.

It is the object of the present investigation to determine what other functions besides the cotangent are thus representable by two Laplace integrals. In order to obtain a series development in terms of exponentials for our generating function $f(z)$, we considered functions which were simply periodic, and to obtain a partial fraction development we took these functions as also meromorphic. By considering the nature and position of the poles in the period strip we have found two classes of functions which admit of the desired representation. As special cases of these are included, besides the cotangent, the other meromorphic trigonometric functions: the tangent, cosecant, and secant, and also the corresponding hyperbolic functions. As special cases of the equalities of the determining series are the linear transformations of the four theta null functions.

In what follows we shall be considering functions simply periodic with primitive period $\gamma = re^{\varphi i}$, $0 \leqq \varphi < \pi$. For convenience of statement we define as the *primitive period-strip* the strip of the z-plane between the straight lines perpendicular to the vectors $z = \gamma/2$ and $z = -\gamma/2$. Moreover, we include *both* these straight lines as part of this primitive strip.

2. A preliminary lemma.

We first prove the

LEMMA. *Let $f(z)$ be a single-valued meromorphic simply periodic function with primitive period γ. Further, let $f(z)$ be bounded at both ends of the primitive period strip and let its poles in this strip be simple and located in the points $a_1, \cdots, a_k$. Then*

$$f(z) = \sum_{i=1}^{k} c_i \frac{\pi}{\gamma}\operatorname{ctn}\frac{\pi}{\gamma}(z - a_i) + C, \tag{7}$$

where c_i is the residue in the pole[5] a_i and C is a suitably chosen constant.

[5] According to our definition of the primitive period strip, c_i is the residue in the pole a_i provided that $a_i \neq \gamma/2$ or $-\gamma/2$, in which cases c_i is half the residue.

Consider the function ctn z. Its partial fraction representation is[6]

$$\operatorname{ctn} z = \frac{1}{z} + \sum_{n=-\infty}^{\infty}{}' \left\{ \frac{1}{z - n\pi} + \frac{1}{n\pi} \right\},$$

the series converging absolutely for all $z \neq n\pi$ and uniformly in any region not including any of the points $n\pi$. Hence

$$c_i \frac{\pi}{\gamma} \operatorname{ctn} \frac{\pi}{\gamma}(z - a_i) = \frac{c_i}{z - a_i} + \sum_{n=-\infty}^{\infty}{}' \left\{ \frac{c_i}{z - (a_i + n\gamma)} + \frac{c_i}{n\gamma} \right\}$$

is a function, single-valued, simply periodic with the period γ, and having simple poles in the points $a_i + n\gamma$ ($n = 0, \pm 1, \pm 2, \cdots$) with residue c_i at each. Then

$$F(z) \equiv \sum_{i=1}^{k} c_i \frac{\pi}{\gamma} \operatorname{ctn} \frac{\pi}{\gamma}(z - a_i)$$

is the partial fraction development of a meromorphic function with poles in the points $a_i + n\gamma$ ($n = 0, \pm 1, \pm 2, \cdots, i = 1, \cdots, k$), and poles in no other points. But these are the only points in which $f(z)$ has poles. It follows then from the Mittag-Leffler theorem[7] that

$$f(z) = F(z) + G(z),$$

where $G(z)$ is an entire function. Now the only singularities of $f(z)$ in the primitive period strip are its poles in the points $a_1, \cdots, a_k$ and these poles are all contained in $F(z)$. Moreover, $f(z)$ and $F(z)$ are bounded at both ends of a period-strip. Consequently, $G(z)$ can have no singularities at all in a period-strip and is bounded at both ends. But since $f(z)$ and $F(z)$ are both simply periodic with the period γ, $G(z)$ must be also. It follows that $G(z)$ is a constant, and the lemma is established.

3. **The first class of functions.** Our first theorem is the following.

THEOREM A. *Let $f(z)$ be a single-valued meromorphic simply periodic function with the primitive period γ and bounded at the ends of a period-strip. Further, let*

$$f(z) - \sum_{i=1}^{N} \frac{2c_i z}{z^2 - p_i^2\gamma^2} \qquad (0 \leqq p_i \leqq \tfrac{1}{2};\ p_i \neq p_j) \tag{8}$$

be analytic in the primitive period-strip. Then, when $\gamma = \alpha$, real,

$$\frac{f(\sqrt{-s}) - C}{-\sqrt{-s}} = \int_0^\infty e^{-st}\varphi_1(t)\,dt = \int_0^\infty e^{-st}\psi_1(t)\,dt \qquad (R(s) > 0), \tag{9}$$

[6] The prime on the summation sign means that the value $n = 0$ is to be omitted.

[7] See, for instance, W. F. Osgood, *Lehrbuch der Funktionentheorie*, 5th ed., 1928, vol. I, p. 565.

where C is a suitably chosen constant, and

$$
(10)\qquad \begin{cases} \varphi_1(t) = 2\displaystyle\sum_{i=1}^{N}\sum_{n=-\infty}^{\infty} c_i e^{-(n+p_i)^2\alpha^2 t} \\[2ex] \psi_1(t) = 2\sqrt{\dfrac{\pi}{\alpha^2 t}}\displaystyle\sum_{i=1}^{N}\sum_{n=-\infty}^{\infty} c_i \cos(2n\pi p_i) e^{-\frac{n^2\pi^2}{\alpha^2 t}}. \end{cases}
$$

COROLLARY A.1. *When $\gamma = i\alpha$, pure imaginary,*

$$
(11)\qquad \frac{f(-\sqrt{s}) - C}{-\sqrt{s}} = \int_0^\infty e^{-st}\varphi_1(t)\,dt = \int_0^\infty e^{-st}\psi_1(t)\,dt \qquad (R(s) > 0).
$$

COROLLARY A.2.

$$
(12)\qquad \sum_{n=-\infty}^{\infty} e^{-(n+p)^2\alpha^2 t} = \sqrt{\frac{\pi}{\alpha^2 t}}\sum_{n=-\infty}^{\infty} \cos(2n\pi p)\, e^{-\frac{n^2\pi^2}{\alpha^2 t}} \qquad (0 \leqq p \leqq \tfrac{1}{2}).
$$

First of all we note that since

$$
\frac{2z}{z^2 - p^2\gamma^2} = \frac{1}{z + p\gamma} + \frac{1}{z - p\gamma},
$$

the condition that (8) be analytic is equivalent to saying that in the primitive period-strip $f(z)$ has only simple poles which occur in pairs (except when $p_i = 0$), with equal residues c_i, in the points $z = \pm p_i\gamma$, symmetric with respect to the origin. When $p_i = 0$, there is a single pole at $z = 0$ with residue $2c_i$.

First let us consider the case when there is only one such pair of poles, and let

$$
f_1(z) \equiv f(z) \text{ when } N = 1, \qquad c_1 = c, \qquad p_1 = p.
$$

Then, by our lemma,

$$
(13)\quad f_1(z) - C \equiv F(z) = c\frac{\pi}{\gamma}\operatorname{ctn}\frac{\pi}{\gamma}(z + p\gamma) + c\frac{\pi}{\gamma}\operatorname{ctn}\frac{\pi}{\gamma}(z - p\gamma).
$$

Using (2), we have

$$
(14)\qquad \begin{aligned} F(z) = c\frac{\pi}{\gamma}\Bigg\{ & \frac{1}{\pi\gamma^{-1}(z + p\gamma)} + \sum_{n=1}^{\infty} \frac{2\pi\gamma^{-1}(z + p\gamma)}{\pi^2\gamma^{-2}(z + p\gamma)^2 - n^2\pi^2} \\ & + \frac{1}{\pi\gamma^{-1}(z - p\gamma)} + \sum_{n=1}^{\infty} \frac{2\pi\gamma^{-1}(z - p\gamma)}{\pi^2\gamma^{-2}(z - p\gamma)^2 - n^2\pi^2}\Bigg\}. \end{aligned}
$$

Moreover, adding these series together term by term, we are able to write $F(z)$ in the form

$$
(15)\quad F(z) = 2c\left\{\frac{z}{z^2 - p^2\gamma^2} + \sum_{n=1}^{\infty}\left(\frac{z}{z^2 - (n-p)^2\gamma^2} + \frac{z}{z^2 - (n+p)^2\gamma^2}\right)\right\}.
$$

Using (3), we may also write

$$F(z) = -i2c\frac{\pi}{\gamma}\left\{1 + 2\sum_{n=1}^{\infty}\cos(2n\pi p)\,e^{\frac{2n\pi}{\gamma}iz}\right\}. \tag{16}$$

Thus, when $\gamma = \alpha$, real,

$$\begin{aligned}\frac{F(\sqrt{-s})}{-\sqrt{-s}} &= 2c\left\{\frac{1}{s + p^2\alpha^2} + \sum_{n=1}^{\infty}\left(\frac{1}{s + (n-p)^2\alpha^2} + \frac{1}{s + (n+p)^2\alpha^2}\right)\right\} \\ &= 2c\frac{\pi}{\alpha}\left\{\frac{1}{\sqrt{s}} + 2\sum_{n=1}^{\infty}\cos(2n\pi p)\frac{e^{-\frac{2n\pi}{\alpha}\sqrt{s}}}{\sqrt{s}}\right\}.\end{aligned} \tag{17}$$

The terms of each of the above series may be expressed as Laplace integrals,[8] and so, if the series of integrals equals the integral of the series, we have the result that

$$\begin{aligned}\frac{F(\sqrt{-s})}{-\sqrt{-s}} &= 2c\int_0^{\infty} e^{-st}\left[e^{-p^2\alpha^2 t} + \sum_{n=1}^{\infty}\left(e^{-(n-p)^2\alpha^2 t} + e^{-(n+p)^2\alpha^2 t}\right)\right]dt \\ &= 2c\frac{\pi}{\alpha}\int_0^{\infty} e^{-st}\left[\frac{1}{\sqrt{\pi t}} + \frac{2}{\sqrt{\pi t}}\sum_{n=1}^{\infty}\cos(2n\pi p)\,e^{-\frac{n^2\pi^2}{\alpha^2 t}}\right]dt\end{aligned} \tag{18}$$

$$(R(s) > 0).$$

In the first of these integrals the term by term integration is permissible within any interval (a, b), $0 < a \leqq t \leqq b < \infty$, since in this interval

$$\sum e^{-[s+(n\mp p)^2\alpha^2]t} \ll \sum e^{-[\sigma+(n\mp p)^2\alpha^2]a}, \qquad \sigma = R(s) > 0,$$

the dominating series being convergent by comparison with $\Sigma\, n^{-2}$. The interval may now be extended to the infinite one $(0, \infty)$, since

$$\sum\int_0^{\infty} |\,e^{-st}\,|\,e^{-(n\mp p)^2\alpha^2 t}\,dt = \sum\frac{1}{\sigma + (n\mp p)^2\,\alpha^2}$$

is convergent for any $\sigma > 0$.[9] An analogous proof establishes the validity of the term by term integration of the second series of (18).

Since our preliminary lemma shows that the function $f(z)$ of the theorem is a linear combination of functions $f_1(z)$, Theorem A is now established. Moreover, when we take $\gamma = i\alpha$, pure imaginary, and construct the function $\dfrac{F(-\sqrt{s})}{-\sqrt{s}}$

[8] See B. O. Peirce, *A Short Table of Integrals*: #493 with $n = 0$, $a = s + (n \mp p)^2\alpha^2$, and #495 with $x = \sqrt{st}$, $a = \dfrac{n\pi}{\alpha}\sqrt{s}$.

[9] For the test used here see E. C. Titchmarsh, *Theory of Functions*, pp. 44–45.

from (15) and (16), we are led to exactly the same two series as in (17). This establishes our first corollary. The second follows immediately from the uniqueness property of Laplace integrals already mentioned.

Essentially it is the equality of the two series of (17) and the fact that they may each be represented by Laplace integrals that gives us the result of our theorem. From the formal appearance of these series it might seem that these results would hold also for an arbitrary complex value of γ. That this is not so follows from the fact that if in (17) we replace the real α by an arbitrary complex γ, we are led to the integral representation

$$\frac{e^{-\frac{2n\pi}{\gamma}\sqrt{s}}}{\sqrt{s}} = \int_0^\infty e^{-st}\,\frac{e^{-\frac{n^2\pi^2}{\gamma^2 t}}}{\sqrt{\pi t}}\,dt$$

which, for $R(s) > 0$, is valid only when $R\left(\frac{n^2\pi^2}{\gamma^2}\,s\right) > 0$,[10] and this condition requires that γ be real.

4. A second class of functions. Our second theorem is

THEOREM B. *Let $f(z)$ be a single-valued meromorphic simply periodic function with the primitive period γ and bounded at the ends of a period-strip. Further, let*

$$f(z) - \sum_{i=1}^{N} \frac{2c_i p_i \gamma}{z^2 - p_i^2\,\gamma^2} \qquad (0 < p_i < \tfrac{1}{2};\ p_i \neq p_j) \tag{19}$$

be analytic in the primitive period-strip. Then when $\gamma = \alpha$, real,

$$-f(\sqrt{-s}) - C = \int_0^\infty e^{-st}\,\varphi_2(t)\,dt = \int_0^\infty e^{-st}\,\psi_2(t)\,dt \quad (R(s) > 0), \tag{20}$$

where C is a suitably chosen constant, and

$$\begin{cases} \varphi_2(t) = 2\,\alpha \displaystyle\sum_{i=1}^{N} \sum_{n=-\infty}^{\infty} c_i(n + p_i)\, e^{-(n+p_i)^2\alpha^2 t} \\[2ex] \psi_2(t) = \dfrac{2}{\alpha^2}\left(\dfrac{\pi}{t}\right)^{3/2} \displaystyle\sum_{i=1}^{N} \sum_{n=-\infty}^{\infty} c_i\, n \sin(2n\pi p_i)\, e^{-\frac{n^2\pi^2}{\alpha^2 t}}. \end{cases} \tag{21}$$

COROLLARY B.1. *When $\gamma = i\alpha$, pure imaginary,*

$$-f(-\sqrt{s}) - C = \int_0^\infty e^{-st}\,\varphi_2(t)\,dt = \int_0^\infty e^{-st}\,\psi_2(t)\,dt \quad (R(s) > 0). \tag{22}$$

COROLLARY B.2.

$$\sum_{n=-\infty}^{\infty} (n + p)e^{-(n+p)^2\alpha^2 t} = \left(\frac{\pi}{\alpha^2 t}\right)^{3/2} \sum_{n=-\infty}^{\infty} n \sin(2n\pi p) e^{-\frac{n^2\pi^2}{\alpha^2 t}} \quad (0 < p < \tfrac{1}{2}). \tag{23}$$

[10] For complex γ and s this is the condition corresponding to $a > 0$ for the integral #495 of Peirce's tables.

Since

$$\frac{2p\gamma}{z^2 - p^2\gamma^2} = \frac{1}{z - p\gamma} - \frac{1}{z + p\gamma},$$

the condition that (19) be analytic is equivalent to saying that in the primitive period-strip $f(z)$ has only simple poles occurring in pairs, one at $z = p_i\gamma$ with residue c_i and the other at $z = -p_i\gamma$ with residue $-c_i$.

First let us consider the case when there is only one such pair of poles and let

$$f_2(z) \equiv f(z) \text{ when } N = 1, \qquad c_1 = c, \qquad p_1 = p.$$

Then, by our lemma,

$$(24)\quad f_2(z) - C \equiv F(z) = c\frac{\pi}{\gamma}\operatorname{ctn}\frac{\pi}{\gamma}(z - p\gamma) - c\frac{\pi}{\gamma}\operatorname{ctn}\frac{\pi}{\gamma}(z + p\gamma).$$

Using (2), we have

$$(25)\quad \begin{aligned} F(z) &= c\frac{\pi}{\gamma}\left\{\frac{1}{\pi\gamma^{-1}(z - p\gamma)} + \sum_{n=1}^{\infty}\frac{2\pi\gamma^{-1}(z - p\gamma)}{\pi^2\gamma^{-2}(z - p\gamma)^2 - n^2\pi^2}\right. \\ &\qquad\left. - \frac{1}{\pi\gamma^{-1}(z + p\gamma)} - \sum_{n=1}^{\infty}\frac{2\pi\gamma^{-1}(z + p\gamma)}{\pi^2\gamma^{-2}(z + p\gamma)^2 - n^2\pi^2}\right\} \\ &= 2c\left\{\frac{p\gamma}{z^2 - p^2\gamma^2} - \sum_{n=1}^{\infty}\left(\frac{(n - p)\gamma}{z^2 - (n - p)^2\gamma^2} - \frac{(n + p)\gamma}{z^2 - (n + p)^2\gamma^2}\right)\right\}. \end{aligned}$$

Hence, when $\gamma = \alpha$,

$$(26)\quad -F(\sqrt{-s}) = 2c\alpha\left\{\frac{p}{s + p^2\alpha^2} - \sum_{n=1}^{\infty}\left(\frac{n - p}{s + (n - p)^2\alpha^2} - \frac{n + p}{s + (n + p)^2\alpha^2}\right)\right\}.$$

We see now that this series differs from those we have been considering in that the numerators of the fractions increase with n. We should like to conclude that it equals a Laplace integral as before, that is, that

$$(27)\quad \begin{aligned} -F(\sqrt{-s}) = 2c\alpha\int_0^{\infty} e^{-st}\left\{pe^{-p^2\alpha^2 t} - \sum_{n=1}^{\infty}[(n - p)e^{-(n-p)^2\alpha^2 t}\right. \\ \left. - (n + p)e^{-(n+p)^2\alpha^2 t}]\right\} dt. \end{aligned}$$

This is indeed true, but the proof does not follow quite as readily as in the previous cases. If we let

$$u_n(t) = (n - p)e^{-(n-p)^2\alpha^2 t} - (n + p)e^{-(n+p)^2\alpha^2 t},$$

then in any interval $0 < a \leqq t \leqq b < \infty$,

$$\sum e^{-st}u_n(t) \ll \sum e^{-\sigma a}[(n - p)e^{-(n-p)^2\alpha^2 a} + (n + p)e^{-(n+p)^2\alpha^2 a}],$$

convergent by comparison with $\sum \frac{1}{n^2}$ for any $\sigma > 0$. Hence

$$\int_a^b \sum e^{-st} u_n(t)\,dt = \sum \int_a^b e^{-st} u_n(t)\,dt. \tag{27a}$$

Now we may write

$$u_n(t) = e^{-(n+p)^2\alpha^2 t}[(n-p)e^{4np\alpha^2 t} - (n+p)].$$

Furthermore, $e^{-(n+p)^2\alpha^2 t} > 0$ for all t, and in the interval $a \leqq t$,

$$(n-p)e^{4np\alpha^2 t} - (n+p) \geqq (n-p)e^{4np\alpha^2 a} - (n+p) > 0$$

for sufficiently large n, $n > N(a)$. Hence,

$$\sum_{n=N}^{\infty} \int_a^{\infty} | e^{-st} u_n(t) |\,dt = \sum_{n=N}^{\infty} \int_a^{\infty} e^{-\sigma t} u_n(t)\,dt$$

$$= \sum_{n=N}^{\infty} \left\{ \frac{(n-p)e^{-[\sigma+(n-p)^2\alpha^2]a}}{\sigma + (n-p)^2\alpha^2} - \frac{(n+p)e^{-[\sigma+(n+p)^2\alpha^2]a}}{\sigma+(n+p)^2\alpha^2} \right\} \qquad (a > 0),$$

and this last series is absolutely convergent by comparison with $\sum \frac{1}{n^2}$. Therefore by the test previously used, we may let b become infinite, and we get

$$\int_a^{\infty} \sum e^{-st} u_n(t)\,dt = \sum \int_a^{\infty} e^{-st} u_n(t)\,dt \qquad (a > 0). \tag{27b}$$

We still need to show that we can let $a \to 0$. Unfortunately $u_n(t)$ does not remain positive for all small t. But consider the series on the right of the above equation. If we replace s by σ, we have a series the terms of which, for a fixed σ, are functions of a, and this series, written in the form

$$\sum_{n=1}^{\infty} \left\{ \int_a^{\infty} e^{-\sigma t}(n-p)e^{-(n-p)^2\alpha^2 t}\,dt - \int_a^{\infty} e^{-\sigma t}(n+p)e^{-(n+p)^2\alpha^2 t}\,dt \right\}$$

$$\equiv \sum_{n=1}^{\infty} \{v_n(a) - w_n(a)\}, \qquad \sigma \text{ fixed},$$

is a convergent alternating series for n sufficiently large. For, first of all, by actually evaluating the integrals $v_n(a)$ and $w_n(a)$, we see that they approach 0 as n becomes infinite, $(a \geqq 0)$. Secondly, we can show that

$$v_n(a) > w_n(a) \text{ and } w_n(a) > v_{n+1}(a).$$

The first of these will be true, provided that

$$\frac{(n-p)e^{-[\sigma+(n-p)^2\alpha^2]a}}{\sigma+(n-p)^2\alpha^2} - \frac{(n+p)e^{-[\sigma+(n+p)^2\alpha^2]a}}{\sigma+(n+p)^2\alpha^2} > 0,$$

and after factoring out $e^{-[\sigma+(n+p)^2\alpha^2]a}$, we find that we need only consider

$$\frac{(n-p)e^{4np\alpha^2 a}}{\sigma+(n-p)^2\alpha^2} - \frac{n+p}{\sigma+(n+p)^2\alpha^2} \geqq \frac{2p[(n^2-p^2)\alpha^2-\sigma]}{[\sigma+(n-p)^2\alpha^2][\sigma+(n+p)^2\alpha^2]},$$

and this is > 0 if $(n^2 - p^2)\alpha^2 > \sigma$. For any fixed σ, there is a definite number M such that for $n > M$, independent of a, this is true. In an exactly similar manner, making use of the fact that $p < \frac{1}{2}$, we can show that for $n > M$, $w_n(a) > v_{n+1}(a)$. This proves our statement that the series under consideration is a convergent alternating series for n sufficiently large. Consequently, the absolute value of the remainder after $2m - 2$ terms is less than the absolute value of the first term of the remainder. That is, if

$$R_m(a) = \sum_{n=m}^{\infty} [v_n(a) - w_n(a)],$$

then, for $m > M$, independent of $a \geqq 0$,

$$| R_m(a) | < \left| \int_a^\infty e^{-\sigma t} v_m(t)\, dt \right| \leqq \int_0^\infty e^{-\sigma t} v_m(t)\, dt = \frac{n-p}{\sigma+(n-p)^2\alpha^2},$$

and this approaches zero as n becomes infinite. Hence, we can find an $M' > M$, independent of a, such that for any given ϵ,

$$| R_m(a) | < \epsilon, \qquad \text{for } m > M',$$

and accordingly the series on the right of (27b) converges uniformly with respect to a for $a \geqq 0$. Hence, this series defines a continuous function for $a \geqq 0$ and the limit of the series equals the series of the limits. Therefore, we have

$$\begin{aligned} \int_0^\infty \sum e^{-\sigma t} u_n(t)\, dt &= \sum \int_0^\infty e^{-\sigma t} u_n(t)\, dt \\ &= \sum \left(\frac{n-p}{\sigma+(n-p)^2\alpha^2} - \frac{n+p}{\sigma+(n+p)^2\alpha^2} \right). \end{aligned} \tag{27c}$$

This is proved for real σ and by analytic continuation it will hold if σ is replaced by any s whose real part is > 0. Hence, the equality (27) is verified.

Further, using (3), we may show that

$$F(z) = -4c\frac{\pi}{\gamma} \sum_{n=1}^{\infty} \sin (2n\pi p) e^{\frac{2n\pi}{\gamma} iz}. \tag{28}$$

Hence, when $\gamma = \alpha$,

$$-F(\sqrt{-s}) = 4c\frac{\pi}{\alpha} \sum_{n=1}^{\infty} \sin (2n\pi p) e^{-\frac{2n\pi}{\alpha}\sqrt{s}}. \tag{29}$$

The terms here are not quite the same as we have had before but there is no essential difference in establishing the desired result that

$$(30) \qquad -F(\sqrt{-s}) = 4c\,\frac{\pi}{\alpha}\int_0^\infty e^{-st}\sum_{n=1}^{\infty}\frac{n\pi}{\alpha}\sin(2n\pi p)\,\frac{e^{-\frac{n^2\pi^2}{\alpha^2 t}}}{t\sqrt{\pi t}}\,dt.$$

Equations (27) and (30) taken in conjunction with the preliminary lemma prove our theorem. The corollaries follow exactly as they did in the case of Theorem A.

5. **Examples.** It has already been mentioned that the cotangent is a special case of this theory, satisfying the conditions of Theorem A. For this function $\gamma = \pi$, $N = 1$, $p_1 = 0$, $c_1 = \frac{1}{2}$. The other meromorphic trigonometric functions are also special cases, the tangent and cosecant coming under Theorem A and the secant under Theorem B. These four trigonometric functions may be grouped as follows:

$$f(z) = \operatorname{ctn} z\colon \gamma = \pi, \quad p_1 = 0, \quad c_1 = \tfrac{1}{2}, \quad (N = 1) - A$$

$$f(z) = \tan z\colon \gamma = \pi, \quad p_1 = \tfrac{1}{2}, \quad c_1 = -1, \quad (N = 1) - A$$

$$f(z) = \csc z\colon \gamma = 2\pi, p_1 = 0, p_2 = \tfrac{1}{2}, c_1 = \tfrac{1}{2}, c_2 = -1, (N = 2) - A$$

$$f(z) = \sec z\colon \gamma = 2\pi, \quad p_1 = \tfrac{1}{4}, \quad c_1 = -1, \quad (N = 1) - B.$$

Then further, the four meromorphic hyperbolic functions are special cases of Corollaries A.1 and B.1, the hyperbolic cotangent, tangent, and cosecant coming under the first of these corollaries and the hyperbolic secant under the second.

Finally, we mention the following example in which the additive constant C of the function $\dfrac{f(-\sqrt{s}) - C}{-\sqrt{s}}$ is not zero as it has been in all the preceding examples:

$$f(z) = \frac{1}{e^z - 1}, \quad C = -\tfrac{1}{2}, \quad \gamma = 2\pi i, \quad p_1 = 0, \quad c_1 = \tfrac{1}{2}, \quad N = 1.$$

This function satisfies the conditions of Corollary A.1.

Moreover, from equations (12) and (23) we may, by suitable choices of p, obtain the linear transformations of the theta null functions. Thus when $p = 0$, equation (12) becomes

$$\sum_{n=-\infty}^{\infty} e^{-n^2\alpha^2 t} = \sqrt{\frac{\pi}{\alpha^2 t}}\sum_{n=-\infty}^{\infty} e^{-\frac{n^2\pi^2}{\alpha^2 t}},$$

which may be written

$$\vartheta_3\left(0, \frac{i\alpha^2 t}{\pi}\right) = \sqrt{\frac{\pi}{\alpha^2 t}}\, \vartheta_3\left(0, -\frac{\pi}{i\alpha^2 t}\right),$$

the linear transformation of the theta null-function ϑ_3 for $\nu = 0$, $\tau = \dfrac{i\alpha^2 t}{\pi}$.[11] When $p = \frac{1}{2}$, (12) becomes

$$\vartheta_2\left(0, \frac{i\alpha^2 t}{\pi}\right) = \sqrt{\frac{\pi}{\alpha^2 t}}\, \vartheta_0\left(0, -\frac{\pi}{i\alpha^2 t}\right),$$

the linear transformation from ϑ_2 to ϑ_0 for $\nu = 0$, $\tau = \dfrac{i\alpha^2 t}{\pi}$. If we take equation (12) for $p = 0$ and $p = \frac{1}{2}$, subtract the second equation from the first, and make certain rearrangements and combinations, we have the linear transformation from ϑ_0 to ϑ_2 for $\nu = 0$, $\tau = \dfrac{i\alpha^2 t}{4\pi}$,

$$\vartheta_0\left(0, \frac{i\alpha^2 t}{4\pi}\right) = \sqrt{\frac{4\pi}{\alpha^2 t}}\, \vartheta_2\left(0, -\frac{4\pi}{i\alpha^2 t}\right).$$

Finally, if in (23) we take $p = \frac{1}{4}$, that equation can be written

$$\vartheta_1'\left(0, \frac{i\alpha^2 t}{4\pi}\right) = \left(\frac{4\pi}{\alpha^2 t}\right)^{3/2} \vartheta_1'\left(0, -\frac{4\pi}{i\alpha^2 t}\right),$$

the linear transformation for $\nu = 0$, $\tau = \dfrac{i\alpha^2 t}{4\pi}$ of the derivative of ϑ_1 with respect to the variable ν.

Georgia School of Technology.

[11] The notation for the Theta functions used here is that found in Hurwitz-Courant, *Funktionentheorie*, 3d ed., vol. III, part II. See particularly Chapters 2 and 7.

A PROBLEM OF ZERMELO IN THE CALCULUS OF VARIATIONS

By William L. Duren, Jr.

1. **Types of relative minima.** Let $y^i(x)$ stand for the i-th derivative of the function $y(x)$ and $y^0(x)$ for $y(x)$ itself. Then an admissible arc E

$$y = y(x) \qquad (x_1 \leqq x \leqq x_2)$$

which joins the points (x_1, y_{01}) and (x_2, y_{02}) will be said to furnish a relative minimum of order r ($r = 0, 1, \cdots, n$) to the problem of Zermelo if there exists a neighborhood N_r of the elements $(x, y, y', \cdots, y^r)$ belonging to E such that E gives to the integral

$$J = \int_{x_1}^{x_2} f(x, y, y', \cdots, y^n)dx \tag{1}$$

a smaller value than that given by every other of a class of admissible arcs C joining the ends of E and having the elements $(x, y, y', \cdots, y^r)$ in N_r. The term admissible arc in this statement will be used in several senses later defined.

The problem of minimizing an integral "with higher derivatives in the integrand" is an old one and is commonly studied as a special problem of Lagrange. However, it will appear that neither necessary nor sufficient conditions for relative minima of orders less than $n - 1$ can be obtained from the general theory of the Lagrange problem in its present form. Thus the classification of these types of relative minima which was first made by Zermelo[1] sets the problem apart and justifies the title of this paper, though Zermelo carried through the analysis for orders n and $n - 1$ only.

In studying relative minima of order r, or weaker ones, one might specify that the elements $(x, y, y', \cdots, y^r)$ are fixed at x_1 and x_2. When this is the case we will speak of the problem of Zermelo with end elements of order r fixed. For the sake of simplicity it will be understood hereafter that unless the contrary is specified we are studying the problem of Zermelo with end elements of order 0 fixed as in the first statement of the problem.

2. **Transformation of the problem into a problem of Lagrange.** If we introduce new variables defined by the equations

$$y_i = y^i(x) \qquad (i = 0, 1, \cdots, n - 1), \tag{2}$$

Received July 7, 1936.

[1] E. Zermelo, *Untersuchungen zur Variationsrechnung*, Dissertation, Berlin, 1894, p. 29.

the problem of Zermelo becomes formally equivalent to the problem of Lagrange with variable end points in which one seeks to minimize the integral

$$\int_{x_1}^{x_2} f(x, y_0, \cdots, y_{n-1}; y'_{n-1})dx$$

in a class of arcs which satisfy the end conditions $y_{01} - y_0(x_1) = y_{02} - y_0(x_2) = 0$ and the differential equations

$$(3) \qquad y'_{\alpha-1} - y_\alpha = 0 \qquad (\alpha = 1, \cdots, n-1).$$

We shall say that an arc (1) for the problem of Zermelo is L-admissible if the equations (2) transform it into an admissible arc in the usual sense[2] for the problem of Lagrange. Clearly an L-admissible arc is of class C^{n-1}. It is seen from equations (2) that weak and strong relative minima for the associated problem of Lagrange correspond respectively to relative minima of orders n and $n-1$ for the problem of Zermelo. If any theorem on relative minima from the general problem of Lagrange is to translate into a theorem on relative minima of order less than $n-1$ for the problem of Zermelo, it must be a theorem in which some of the variables are unrestricted. No such theorem exists.

On account of the form of the equations (3) every L-admissible subarc is normal. The function F which occurs in the theory of the Lagrange problem can then be written in the alternative forms

$$(4) \qquad \begin{aligned} F(x, y, y', \lambda) &= f(x, y_0, \cdots, y_{n-1}; y'_{n-1}) + \sum_\alpha \lambda_\alpha(y'_\alpha - y_\alpha), \\ F_1(x, y, y', \mu) &= f(x, y_0; y'_0, \cdots, y'_{n-1}) + \sum_\alpha \mu_\alpha(y'_\alpha - y_\alpha). \end{aligned}$$

The non-tangency condition is also fulfilled automatically.

The following is a summary of some of the results on the problem of Zermelo which may be obtained by translating the theory of the general problem of Lagrange by means of equations (3).

I. *Along an L-admissible arc E which furnishes a relative minimum of order n to the problem of Zermelo the equation*

$$(5) \qquad \sum_{k=0}^{n} (-1)^k \int_{x_1}^{x} \overset{k}{\cdots} \int_{x_1}^{x} f_{y^{n-k}}(dx)^k = \sum_{i=0}^{n-1} c_i(x - x_1)^i$$

holds identically. In it the symbol $\int_{x_1}^{x} \overset{k}{\cdots} \int_{x_1}^{x}$ *stands for a k-tuple integral and the quantities* $c_0, \cdots, c_{n-1}$ *are suitably chosen constants.*

If E is an arc of class C^{2n} and if the integrand function f is of class C^{n+1} in the arguments $(x, y^0, \cdots, y^n)$, the quantities

$$(6) \qquad L_k = \sum_{i=0}^{n-k} \left(-\frac{d}{dx}\right)^i f_{y^{k+i}} \qquad (k = 0, 1, \cdots, n)$$

[2] G. A. Bliss, Amer. Journal of Math., vol. 52 (1930), p. 677.

can be calculated on E. Along such an arc the Euler equation (5) can be written in the form

$$L_0 \equiv 0, \tag{5'}$$

and the multipliers have the values $\lambda_\alpha = L_\alpha$ $(\alpha = 1, \cdots, n-1)$. An L-admissible arc of class C^{2n} which satisfies equation (5′) is called an *extremal* for the problem of Zermelo. If $f_{y^n y^n} \neq 0$ along an extremal, it is said to be non-singular. Furthermore a minimizing extremal satisfies the following transversality condition.

At both end elements of an L-admissible extremal which affords a relative minimum of order n to the problem of Zermelo with fixed end elements of order k, the equations

$$L_s = 0 \qquad (s = k+2, \cdots, n) \tag{7}$$

must hold.

It is to be understood that the conditions (7) are automatically satisfied in problems with fixed end elements of order $k \geqq n - 1$. Incidentally, these transversality conditions imply that the constants $c_0, \cdots, c_{n-k-2}$ in (5) are zero.

II. *At every element $(x, y^0, \cdots, y^n)$ of an L-admissible arc which satisfies the first necessary condition* (5) *and furnishes to the problem of Zermelo a relative minimum of order $n - 1$ the inequality*

$$f(x, y^0, \cdots, y^{n-1}; Y^n) - f(x, y^0, \cdots, y^n) - (Y^n - y^n) f_{y^n}(x, y^0, \cdots, y^n) \geqq 0$$

must hold for every L-admissible set $(x, y^0, \cdots, y^{n-1}; Y^n) \neq (x, y^0, \cdots, y^n)$.

III. *The elements of an L-admissible arc which satisfies* (5) *and furnishes a relative minimum of order n satisfy the condition*

$$f_{y^n y^n} \geqq 0.$$

The accessory minimum problem for a non-singular extremal is formally equivalent to the Lagrange problem of minimizing the integral

$$\int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx,$$

where

$$2\omega = \sum_{\mu,\nu=0}^{n-1} f_{y^\mu y^\nu} \eta_\mu \eta_\nu + 2 \sum_{\mu=0}^{n-1} f_{y^\mu y^n} \eta_\mu \eta'_{n-1} + f_{y^n y^n} \eta'^2_{n-1}$$

in a class of admissible arcs which satisfy the equations

$$\Phi_\alpha = \eta'_{\alpha-1} - \eta_\alpha = 0 \qquad (\alpha = 1, \cdots, n-1)$$

and the end conditions $\eta_0(x_1) = \eta_0(x_2) = 0$. In terms of the canonical variables (x, η, ζ) and the hamiltonian function $H(x, \eta, \zeta)$, the canonical accessory equations are

$$\eta'_i = H_{\zeta_i}, \qquad \zeta'_i = -H_{\eta_i} \qquad (i = 0, \cdots, n-1), \tag{8}$$

and the secondary transversality and end conditions are

$$\zeta_s(x_1) = \eta_0(x_1) = 0, \tag{9}$$

$$\zeta_s(x_2) = \eta_0(x_2) = 0 \qquad (s = 1, \cdots, n-1). \tag{10}$$

The conditions (9) and (10) define respectively the conjugate systems η_{ij}, ζ_{ij} and u_{ik}, v_{ik} of solutions of the canonical accessory equations (8). In terms of these notations we may state the following necessary condition:[3]

IV. *Along a non-singular extremal which furnishes a relative minimum of order n to the problem of Zermelo, with respect to L-admissible arcs, the relation*

$$\sum_{i,j,k=0}^{n-1} (\zeta_{ij}u_{ik} - \eta_{ij}v_{ik})\, a_j b_k \geqq 0$$

holds for every set x, a_j, b_k for which x is on the interval (x_1, x_2) and for which the equation

$$\sum_{j=0}^{n-1} \eta_{ij}(x)\, a_j = \sum_{k=0}^{n-1} u_{ik}(x)\, b_k$$

is satisfied.

An L-admissible extremal arc E which satisfies the transversality conditions[4] (7) with $k = 0$ and the strengthened conditions[4] II_N', III′, IV′ furnishes to the problem of Zermelo a relative minimum of order $n - 1$ in a class of L-admissible arcs.

3. **Extended admissibility.** We now extend the definition of admissibility and say that an arc C is L_k-admissible if it is composed of a finite number of L-admissible arcs of class C^n and is such that the functions $y(x), y^1(x), \cdots, y^k(x)$ belonging to it are continuous on the interval (x_1, x_2).

Let E be an L-admissible extremal arc which furnishes to the problem of Zermelo a relative minimum of order n with respect to neighboring L_k-admissible arcs. Let 3 be a point on such an arc between 1 and 2. Then the arc E_{32} must furnish a minimum with respect to neighboring L-admissible arcs which join the same end elements $(x, y, y', \cdots, y^k)$. Hence the transversality condition (7) must be satisfied at 3. From the form of the functions $L_{k+2}, \cdots, L_n$ given in (6) it is easily seen that a necessary and sufficient condition for their vanishing identically on E is that the functions $f_{y^{k+2}}, \cdots, f_{y^n}$ vanish identically on E. Hence we have proved the following stronger first necessary condition.

I_k. *At every element of an L-admissible arc of class C^{n+k+1} which furnishes a relative minimum of order n to the problem of Zermelo with respect to neighboring L_k-admissible arcs, the differential equations*

[3] M. R. Hestenes, Trans. Am. Math. Soc., vol. 36 (1934), pp. 793–818.

[4] M. R. Hestenes, loc. cit., p. 815.

$$(11)\qquad \sum_{i=0}^{k+1}\left(-\frac{d}{dx}\right)^i f_{y^i} = 0,$$

$$f_{y^s} = 0 \qquad (s = k+2, \cdots, n)$$

must hold.[5]

Now let E be an L_k-admissible arc along which the equations (13) hold. We make a simple extension of the Weierstrass construction as follows. We choose an arbitrary L_k-admissible arc C: $y = Y(x)$ which has the element $(x, Y, Y', \cdots, Y^k)$ in common with E at 3. Then through a neighboring point 4 of E we construct a family of L_k-admissible arcs D_{54} having contact of order k with E at 4 and with C at a point 5 of C which is near 3. If $x_3 \leqq x_5 < x_4$, the composite arcs $E_{13} + C_{35} + D_{54} + E_{42}$ are L_k-admissible and the integral J taken over such a composite arc is a function of x_5 whose derivative $J'(x_3 + 0)$ must not be negative if E is to be a minimizing arc in the family of composite arcs. Calculating this derivative with the aid of (11), we are led to the following condition.

II$_r$. *If an L-admissible arc of class C^{n+r+1} satisfies the condition* I$_r$ *and furnishes a relative minimum of order r to the problem of Zermelo with respect to neighboring L_r-admissible arcs, then the inequality*

$$(12)\qquad \mathcal{E}_r(x, y^0, y', \cdots, y^n; Y^{r+1}, \cdots, Y^n) =$$
$$f(x, y^0, \cdots, y^r; Y^{r+1}, \cdots, Y^n) - f(x, y^0, \cdots, y^n) -$$
$$(Y^{r+1} - y^{r+1}) f_{y^{r+1}}(x, y^0, \cdots, y^n) \geqq 0$$

must hold at every element $(x, y^0, \cdots, y^n)$ of E and for every admissible set $(x, y^0, \cdots, y^r, Y^{r+1}, \cdots, Y^n) \neq (x, y^0, \cdots, y^n)$.

This $\mathcal{E}$-function, considered as a function of $Y^{r+1}, \cdots, Y^n$ must have a relative minimum for $(Y^{r+1}, \cdots, Y^n) = (y^{r+1}, \cdots, y^n)$ if the arc E is to furnish a relative minimum of order n to the problem of Zermelo with respect to neighboring L_r-admissible arcs. The first partial derivatives of this function vanish at $(Y^{r+1}, \cdots, Y^n) = (y^{r+1}, \cdots, y^n)$ as is seen with the aid of the equations (11). The necessary condition on the second derivatives leads us to another analogue of the Legendre condition.[6]

III$_k$. *If the L-admissible arc of class C^{n+k+1} satisfies the condition* I$_k$ *and furnishes to the problem of Zermelo a relative minimum of order n with respect to neighboring L_k-admissible arcs, then the inequality*

$$\sum_{\mu,\nu=k+1}^{n} f_{y^\mu y^\nu}(x, y^0, \cdots, y^n) z_\mu z_\nu \geqq 0$$

holds for all elements $(x, y^0, \cdots, y^n)$ of E and all sets of numbers $(z_{k+1}, \cdots, z_n)$.

Let E be a non-singular extremal arc satisfying the condition I$_k$ and let $y =$

[5] This condition for the case $n = 2$ was proved by H. H. Pixley, *Contributions to the Calculus of Variations*, 1931–1932, The University of Chicago, pp. 133–189.

[6] Pixley, loc. cit., p. 163.

$y(x, b)$ be a family of L_k-admissible arcs which contains E for $b = 0$. The integral J taken over the arcs of this family defines a function $J(b)$. On account of the condition I_k it is possible to calculate the second variation along E and it is found to have the same value as in §2, namely

$$J''(0) = \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx,$$

where 2ω is the quadratic form occurring in the statement of the condition IV of §2. The accessory minimum problem becomes that of minimizing the integral $J''(0)$ in a class of L_k-admissible arcs

$$\eta_i = \eta_i(x) \qquad (i = 0, 1, \cdots, n-1)$$

which satisfy the equations

$$\eta'_{\alpha-1} - \eta_\alpha = 0 \qquad (\alpha = 1, \cdots, n-1)$$

and the end conditions $\eta_0(x_1) = \eta_0(x_2) = 0$. Again we make an argument along similar lines to that of Hestenes[7] and consider the two conjugate systems, η_{ij}, ζ_{ij} and u_{il}, v_{il}, of solutions of the canonical accessory equations (8), determined respectively by the end conditions (9) and (10). A composite arc defined by $\eta_i(x)$ on the interval (x_1, x_3) and $u_i(x)$ on the interval (x_3, x_2), where

$$\eta_i = \sum_{j=0}^{n-1} \eta_{ij} a_j, \qquad u_i = \sum_{l=0}^{n-1} u_{il} b_l \tag{13}$$

is L_k-admissible, provided that at $x = x_3$ the equations

$$\sum_{j=0}^{n-1} \eta_{\sigma j}(x_3) a_j = \sum_{l=0}^{n-1} u_{\sigma l}(x_3) b_l \qquad (\sigma = 0, 1, \cdots, k) \tag{14}$$

are satisfied. Hestenes has shown that along such an arc the second variation has the form

$$J''(0) = \sum_{i=0}^{n-1} (\eta_i(x_3)\zeta_i(x_3) - u_i(x_3)v_i(x_3)). \tag{15}$$

With these notations we can state the following condition of Jacobi type.

IV_k. *Let E be a non-singular extremal, satisfying the condition* I_k *and the transversality conditions $L_2 = \cdots = L_n = 0$, and furnishing to the problem of Zermelo a relative minimum of order n with respect to neighboring L_k-admissible arcs. Then the inequality $J''(0) \geqq 0$ must hold for arbitrary sets x, a_j, b_l with x interior to (x_1, x_2) subject to* (13) *and* (14), *where $J''(0)$ has the form* (15).

More especially we have the following corollaries.

COROLLARY 1. *For the arc E of condition* IV_k *to furnish a relative minimum of order n with respect to neighboring L_k-admissible arcs it is necessary that the inequality*

$$\sum_{i,j,l=0}^{n-1} (\zeta_{ij} u_{il} - \eta_{ij} v_{il}) a_j b_l \geqq 0$$

[7] Hestenes, loc. cit., p. 801.

hold for all sets x, a_j, b_l satisfying (13) *and* (14) *and the further condition*

$$\sum_{j=0}^{n-1} \zeta_{sj}(x)a_j = -\sum_{l=0}^{n-1} v_{sl}(x)b_l \qquad (s = k+1, \cdots, n-1). \tag{16}$$

This follows from the condition IV_k, since the equations (14) imply that at x_3, $\eta_\sigma(x_3) = u_\sigma(x_3)$, while equations (16) imply that $\zeta_s(x_3) = -v_s(x_3)$. Thus (15) takes the form

$$J''(0) = \sum_{i=0}^{n-1} (\zeta_i(x_3)u_i(x_3) - \eta_i(x_3)v_i(x_3))$$

and the corollary follows from (13).

COROLLARY 2. *If the equations* (16) *in Corollary* 1 *are replaced by*

$$\sum_{j=0}^{n-1} \zeta_{sj}(x)a_j = \sum_{l=0}^{n-1} v_{sl}(x)b_l = 0, \tag{16'}$$

we have the necessary condition

$$\sum_{\sigma=0}^{k} \sum_{j,l=0}^{n-1} (\zeta_{\sigma j}u_{\sigma l} - \eta_{\sigma j}v_{\sigma l})a_j b_l \geqq 0$$

for the relative minimum of order n.

4. **Fields.** We consider a one-parameter family of L_k-admissible extremal arcs which satisfy the condition I_k. Let the ends of these arcs describe two L_k-admissible arcs C and D. If we return to the representation of the problem of Zermelo as a problem of Lagrange as described in §2, the condition I_k implies that

$$F_{y_s'} = 0 \qquad (s = k+1, \cdots, n-1)$$

hold identically on each member of the family. These equations make it possible to calculate the differential of the integral J taken along an extremal of such a family of arcs. In fact, one finds

$$dJ = \left(F - \sum_{\sigma=0}^{k} y_\sigma' F_{y_\sigma'}\right)dx + \sum_{\sigma=0}^{k} F_{y_\sigma'}\,dy_\sigma \Big|_1^2$$

in which the arguments of F are the x, y_i, y_i', λ belonging to the extremal.[8]

Thus the invariant integral J_k^* has the differentials $dy_{k+2}, \cdots, dy_n$ missing and may be evaluated on L_k-admissible arcs. With it the auxiliary formula

$$J(E_{56}) - J(E_{34}) = J_k^*(D_{46}) - J_k^*(C_{35})$$

is valid, where E_{56} and E_{34} are two of the extremals of the family. This invariant integral suggests the following.

[8] G. A. Bliss, Amer. Journ. of Math., vol. 52 (1930), p. 714.

Definition of a field of type k. A field of type k is a region $\mathfrak{F}_k$ *of* $x, y_0, \cdots, y_k$*-space with which is associated a set of functions*

$$p_i(x, y_0, \cdots, y_k), \quad l_\alpha(x, y_0, \cdots, y_k) \quad (i = 1, \cdots, n; \alpha = 1, \cdots, n-1)$$

of class C′ in $\mathfrak{F}_k$*, which are such that*

(a) *the sets* $(x, y, y', \cdots, y^n) = (x, y_0, p_1, \cdots, p_n)$ *are L-admissible,*

(b) *the line integral*

$$J_k^* = \int (F - \sum_\sigma y'_\sigma F_{y'_\sigma})dx + \sum_\sigma F_{y'_\sigma}\, dy_\sigma$$

formed with the arguments $(x, y_0, \cdots, y_k, p_{k+1}, \cdots, p_{n-1}; p_1, \cdots, p_n; l_1, \cdots, l_{n-1})$ *is independent of the path in* $\mathfrak{F}_k$.

For example, we take a field of type 0. In the notations of the problem of Zermelo we suppress the multipliers and find that such a field is a region $\mathfrak{F}_0$ of x, y-space with which is associated a set of functions $p_1(x, y), \cdots, p_n(x, y)$ such that the integral

$$J_0^* = \int (f - p_1 f_{y'})dx + f_{y'}\, dy$$

is independent of the path in $\mathfrak{F}_0$. The arguments of f and $f_{y'}$ are $(x, y, p_1, \cdots, p_n) = (x, y, p_1, \partial p_1/\partial x, \cdots, \partial^{n-1} p_1/\partial x^{n-1})$.

Every non-singular extremal arc E for the associated Lagrange problem to the problem of Zermelo having a conjugate system U_{ij}, V_{ij} of solutions of the accessory equations for E can be imbedded in an n-parameter family of extremals

$$(17) \quad y_i = Y_i(x, \alpha_0, \cdots, \alpha_{n-1}), \quad z_i = Z_i(x, \alpha_0, \cdots, \alpha_{n-1}) \quad (i = 0, 1, \cdots, n-1)$$

which contains E for $(\alpha) = (0)$ and $x_1 \leqq x \leqq x_2$ and the variations of the family along E are

$$Y_{i\alpha_j}(x, 0) = U_{ij}(x), \qquad Z_{i\alpha_j}(x, 0) = V_{ij}(x).$$

Now if E satisfies the conditions I_k and if the determinant

$$(18) \qquad \begin{vmatrix} U_{\sigma j} \\ V_{sj} \end{vmatrix} \qquad \begin{pmatrix} \sigma = 0, \cdots, k \\ s = k+1, \cdots, n-1 \\ j = 0, \cdots, n-1 \end{pmatrix}$$

is different from zero along E, the equations

$$y_\sigma = Y_\sigma(x, \alpha)$$

$$0 = Z_s(x, \alpha)$$

have the initial solutions $(\alpha) = (0)$, $y_\sigma = y_\sigma(x)$, $(x_1 \leqq x \leqq x_2)$, belonging to E and consequently they have unique solutions $\alpha_i(x, y_0, \cdots, y_k)$ of class C' which vanish on E. These functions, together with the multipliers $\Lambda_\beta(x, \alpha)$ of the family (17) determine the functions

$$
(19) \qquad \begin{aligned} p_i(x, y_0, \cdots, y_k) &= Y_{ix}(x, \alpha(x, y_0, \cdots, y_k)) \\ l_\beta(x, y_0, \cdots, y_k) &= \Lambda_\beta(x, \alpha(x, y_0, \cdots, y_k)) \end{aligned}
$$

which are defined and of class C' on a neighborhood $\mathfrak{F}_k$ of the sets $(x, y_0, \cdots, y_k)$ belonging to E. Furthermore the Hilbert integral J^* for the general Lagrange problem when formed with the functions (19) becomes identical with our degenerate integral J_k^* since the functions $F_{y_s'} = Z_s(x, \alpha(x, y_0, \cdots, y_k))$ vanish identically in $\mathfrak{F}_k$. Hence we may use the arguments of the general problem of Lagrange[9] to conclude that J_k^* is independent of the path in $\mathfrak{F}_k$. Thus we have constructed a field of type k about E.

It is noted that all of the extremals of this field, defined by the differential equations

$$
\frac{dy_i}{dx} = p_i(x, y_0, \cdots, y_k),
$$

satisfy the conditions I_k. In the case $k = 0$ our field is determined by a one-parameter family of extremals all of which satisfy the differential equations

$$
f_{y^s}(x, y^0, \cdots, y^n) = 0 \qquad (s = 2, \cdots, n).
$$

In order to establish a lemma on the possibility of imbedding a given extremal in a field of type k we employ a strengthened condition IV_k' which asserts that the condition of Corollary 1 to the condition IV_k holds along E and furthermore that the end and transversality conditions are not conjugate.

LEMMA. *If E is an L-admissible, non-singular extremal arc which satisfies* I_k *and* IV_k', *then there exists a conjugate system $U_{ij}(x)$, $V_{ij}(x)$ of solutions of the canonical accessory equations with a determinant* (18) *which does not vanish on the interval* (x_1, x_2). *Furthermore E can be imbedded in a field of type k.*

Proof. The two conjugate bases of the necessary condition can be replaced[10] by two bases for which

$$
\sum_{i=0}^{n-1} (\zeta_{ij} u_{il} - \eta_{ij} v_{il}) = \delta_{jl}.
$$

The system U_{ij}, V_{ij} defined by the equations

$$
\begin{aligned} U_{\sigma j} &= \eta_{\sigma j} + u_{\sigma j}, \qquad & V_{\sigma j} &= \zeta_{\sigma j} + v_{\sigma j}, \\ U_{sj} &= \eta_{sj} - u_{sj}, & V_{sj} &= \zeta_{sj} - v_{sj} \end{aligned}
$$

is a conjugate system. If the determinant (18) vanishes at x_3, there exists a set of constants a_j, not all zero, with which

$$
\sum_{j=0}^{n-1} U_{\sigma j}(x_3) a_j = \sum_{j=0}^{n-1} [\eta_{\sigma j}(x_3) a_j + u_{\sigma j}(x_3) a_j] = 0,
$$

$$
\sum_{j=0}^{n-1} V_{sj}(x_3) a_j = \sum_{j=0}^{n-1} [\zeta_{sj}(x_3) a_j - v_{sj}(x_3) a_j] = 0.
$$

[9] Bliss, loc. cit., p. 733.
[10] Hestenes, loc. cit., p. 807.

The set $(a_j, b_k) = (a_j, -a_k)$ would satisfy (14) and (16) and would give $J''(0)$ the negative value $\sum_j - a_j a_j$ which contradicts the property IV_k. The second statement in the lemma has already been proved.

5. **Sufficiency theorems.** If E is an L-admissible extremal of a field $\mathcal{F}_k$ of type k which joins two points 1 and 2 in the sense that $y_{01} - y_0(x_1) = y_{02} - y_0(x_2)$, and if C is an L_k-admissible arc which lies in $\mathcal{F}_k$ and joins the same points we can calculate the difference

$$J(C) - J(E) = J(C) - J_k^*(E) = J(C) - J_k^*(C).$$

From the last form we find that

$$J(C) - J(E) = \int_{x_1}^{x_2} [f(x, y_0, \cdots, y_{n-1}, y_{n-1}') \\ - f(x, y_0, \cdots, y_k, p_{k+1}, \cdots, p_n) - l_{k+1}(y_{k+1} - p_{k+1})] dx,$$

where l_{k+1} is evaluated for the arguments $(x, y_0, \cdots, y_k, p_{k+1}, \cdots, p_n)$. In terms of the original notation this formula may be written

$$(20) \quad J(C) - J(E) = \int_{x_1}^{x_2} \mathcal{E}_k(x, y^0, \cdots, y^k, p_{k+1}, \cdots, p_n; y^{k+1}, \cdots, y^n)\, dx,$$

where the functions $(x, y^0, \cdots, y^n)$ belong to C, the arguments of $p_{k+1}, \cdots, p_n$ are $(x, y^0, \cdots, y^k)$, and where $\mathcal{E}_k$ is defined as in (12).

Formula (20) at once justifies the following extension of the usual fundamental sufficiency theorem.

SUFFICIENCY THEOREM 1. *An L-admissible extremal E of a field $\mathcal{F}_k$ of type k furnishes a minimum to the integral with respect to L_k-admissible arcs C which lie in $\mathcal{F}_k$ and join the same end points (x, y_{01}) and (x_2, y_{02}), provided the inequality* (12) *holds for all elements $(x, y^0, \cdots, y^k)$ of $\mathcal{F}_k$.*

We will say that the arc E satisfies the condition II_{rN}' if there exists a neighborhood N of the elements $(x, y^0, \cdots, y^n)$ on E in which the inequality

$$\mathcal{E}_r(x, y^0, \cdots, y^n; Y^{r+1}, \cdots, Y^n) > 0$$

is true for every L_r-admissible set

$$(x, y^0, \cdots, y^r, Y^{r+1}, \cdots, Y^n) \neq (x, y^0, \cdots, y^n).$$

The preceding theorem, together with the continuity properties of the functions p_i, and the lemma of the preceding section give us another theorem.

SUFFICIENCY THEOREM 2. *If E is an L-admissible extremal which satisfies the conditions I_r, II_{rN}', III', IV_r', then E furnishes a proper relative minimum of order r to the problem of Zermelo with respect to L_r-admissible arcs.*

Since the theorem is true, à fortiori, for L-admissible arcs we may translate the statements of the theorem into the language of the associated Lagrange problem of §2. We then have the

COROLLARY. *The transforms of the conditions of the theorem are sufficient to insure that the L-admissible extremal arc E furnishes to the integral*

$$J = \int_{x_1}^{x_2} f(x, y_0, y_1, \cdots, y_{n-1}; y'_{n-1})\,dx$$

of the associated Lagrange problem a minimum with respect to L-admissible arcs which satisfy the same end conditions, $y_{01} - y_0(x_1) = y_{02} - y_0(x_2) = 0$, *and have the sets* $x, y_0(x), \cdots, y_r(x)$ *in a sufficiently small neighborhood of those of E.*

We note that this is a sufficiency condition for a special problem of Lagrange in which not only the derivatives y_i' of comparison arcs but also some of the functions y_i themselves are unrestricted.

6. **Concerning necessary conditions in a class of L-admissible arcs.** There remains the question of finding an extension of the Weierstrass condition which will be a necessary condition that an extremal arc E furnish to the integral (1) a relative minimum of order 0 with respect to neighboring L-admissible arcs, that is, admissible arcs of class C^{n-1}.

For simplicity we consider relative minima of order 0 only. If we followed the ordinary procedure, we would set up a family of L-admissible arcs

$$y = y(x, a) \tag{21}$$

which contains a particular extremal E for $a = 0$ and which satisfies the conditions

$$\lim_{a=0} \frac{\partial^i}{\partial x^i} y(x, a) = y^i(x)$$

for every $x \neq x_3$ on (x_1, x_2) and

$$\lim_{a=0} \frac{\partial^i}{\partial x^i} y(x_3, a) = Y^i,$$

where $(x_3, y(x_3), Y', \cdots, Y^n)$ is an arbitrary L-admissible set. If $n > 1$ no such family exists having the second derivative $y_{xx}(x, a)$ bounded for all values of (x, a) such that $x_1 \leqq x \leqq x_2$, $|a| \leqq \epsilon$. Consequently, in order to construct such a family, it must be assumed that in the class of L-admissible arcs which defines the problem of arguments $y^2, \cdots, y^n$ are unrestricted. When we form the function $J(a)$ by evaluating the integral (1) along the arc of the family (21) having parameter value a, we cannot say that $J(a)$ is continuous at $a = 0$. Thus we could not validate the calculation of the derivative $J'(0)$ which would be expected to give rise to the $\mathcal{E}$-function at 3.

An example will show that the condition II_0 of §3 is not necessary for a relative minimum of order 0 in a class of L-admissible arcs. Consider the problem of minimizing the integral

$$\int_0^\pi (4y''^2 - y'^2)\,dx \tag{22}$$

in the class of all arcs $y = y(x)$ which join the points $(0, 0)$ and $(\pi, 0)$ and have functions $y(x)$ of class C' while the second derivatives $y''(x)$ exist and are continuous on $(0, \pi)$ except at a finite number of points. Thus the class is a class of L_1-admissible arcs. A well known inequality[11] states that

$$\int_0^\pi z^2 dx \leqq 4 \int_0^\pi z'^2 dx$$

holds for all absolutely continuous functions $z(x)$ which vanish at 0 when the second integral exists. Hence in the class of arcs defined for the integral (22) we have

$$\int_0^\pi (y' - y'(0))^2 dx \leqq 4 \int_0^\pi y''^2 dx.$$

Since $y(x)$ vanishes at both end points, one finds that

$$\pi\, y'^2(0) \leqq \int_0^\pi (4y''^2 - y'^2) dx.$$

The equality holds only when $y(x) \equiv 0$ on the interval $(0, \pi)$. The arc $y(x) \equiv 0$ is an extremal which joins the end points $(0, 0)$ and $(\pi, 0)$ and has $f_{y''} \equiv 0$ along it. (I.e., the extremal satisfies the condition I_0 of §3.) This extremal furnishes a relative minimum of order 0, in fact an absolute minimum, to the problem. But along this extremal the $\mathcal{E}$-function of §3 has the value

$$\mathcal{E}_0(x, y, y', y''; Y', Y'') = 4Y''^2 - Y'^2.$$

Since this is not a positive form, we conclude that the condition II_0 of §3 is not necessary for a relative minimum of order 0 in a class of L-admissible arcs, even on an extremal which satisfies the condition I_0. Hence also the condition III_0 is not necessary.

On the other hand, simple examples can be constructed to show that a sufficiency theory for relative minima of order 0 cannot be built upon the Weierstrass condition which comes from the associated problem of Lagrange.

Tulane University.

[11] L. Tonelli, *Fondamenti di Calcolo delle Variazioni*, Bologna, 1929, vol. 2, p. 439. Picard, *Traité d'Analyse*, vol. 3, 1st ed., 1896, p. 115.

NON-SEPARATING TRANSFORMATIONS

By James F. Wardwell

1. **Introduction.** If A is a compact continuum and $T(A) = B$ is a single-valued continuous transformation, then T will be said to be *non-separating* provided that no set $T^{-1}(b)$, $b \in B$, separates A. It is obvious that any non-separating transformation is non-alternating.[1] However, it can easily be seen by simple examples that not every non-alternating transformation is non-separating; not every non-separating transformation is monotone;[2] and not every monotone transformation is non-separating.

Since any continuous transformation between two compact metric spaces A and B is equivalent[3] to an upper semi-continuous decomposition[4] of A into disjoint closed sets where the hyperspace of the decomposition is homeomorphic with B, any non-separating transformation $T(A) = B$ is equivalent to an upper semi-continuous decomposition of A into sets which do not separate A.

All transformations used in this paper will be assumed to be single-valued and continuous.

2. **Some characteristic properties.**

Theorem 2.1. *If A and B are compact continua, a necessary and sufficient condition in order that $T(A) = B$ be non-separating is that T be non-alternating and B contain no cut points.*

Proof. To prove the necessity, in view of our remarks in the above section, we need only show that B can contain no cut points. If, for some point b of B, there were a separation $B - b = B_1 + B_2$, then $T^{-1}(b)$ would separate A into the two mutually separated sets $T^{-1}(B_1)$ and $T^{-1}(B_2)$ because of the continuity of T. The sufficiency follows at once from a theorem of G. T. Whyburn's[5] which states that if B is connected and $T(A) = B$ is non-alternating, then a point x of B is a cut point of B if and only if $T^{-1}(x)$ separates A.

Received August 8, 1936; presented to the American Mathematical Society, April 10, 1936.

[1] A continuous transformation $T(A) = B$ is *non-alternating* provided that for any $x, y \in B$, $T^{-1}(x)$ does not separate $T^{-1}(y)$ in A. See G. T. Whyburn, American Journal of Mathematics, vol. 56 (1934), no. 2, pp. 294–302.

[2] A continuous transformation $T(A) = B$ is *monotone* provided that each set $T^{-1}(b)$, $b \in B$, is connected. See C. B. Morrey, Jr., American Journal of Mathematics, vol. 57 (1935), pp. 17–50, and G. T. Whyburn, loc. cit.

[3] C. Kuratowski, Fundamenta Mathematicae, vol. 11 (1928), pp. 169–185.

[4] R. L. Moore, Transactions of the American Mathematical Society, vol. 27 (1925), pp. 416–428.

[5] See p. 295 of his paper *Non-alternating transformations*, loc. cit.

COROLLARY. *A non-alternating transformation $T(A) = B$ is non-separating if and only if B contains no cut points.*

THEOREM 2.2. *If A and B are compact continua and $T(A) = B$ is non-separating, then for any cut point p of A and for any possible separation $A - p = A_1 + A_2$, $A_i \subset T^{-1}T(p)$ for i equal either 1 or 2. Furthermore $A_j - A_j \cdot T^{-1}T(p)$ is connected for $j = 1$ or 2 and $j \neq i$.*

Proof. Take any cut point p of A and any separation $A - p = A_1 + A_2$. Now, if we let $T(p) = b$,

$$A - T^{-1}(b) = [A_1 - A_1 \cdot T^{-1}(b)] + [A_2 - A_2 \cdot T^{-1}(b)].$$

Hence one of these sets, say $A_1 - A_1 \cdot T^{-1}(b)$, must be vacuous, since T is non-separating. Therefore $A_1 \subset T^{-1}(b)$. Furthermore $A_2 - A_2 \cdot T^{-1}(b)$ must be connected, since we can not have any separation of $A - T^{-1}(b)$.

3. **Product and factor theorems.** Let A be a compact continuum and let $T(A) = B$ be expressed as the product of the transformations $T_1(A) = A'$ and $T_2(A') = B$. That is, we have $T(A) = T_2T_1(A) = T_2(A') = B$.

THEOREM 3.1. *If T is non-separating, T_2 is non-separating regardless of T_1.*

Proof. Let us suppose that T_2 is not non-separating. Then there is some point p of B so that there is a separation $A' - T_2^{-1}(p) = A_1 + A_2$. From this, because of the continuity of T_1, we have the separation $A - T_1^{-1}T_2^{-1}(p) = T_1^{-1}(A_1) + T_1^{-1}(A_2)$. But $T_1^{-1}T_2^{-1}(p) = T^{-1}(p)$. Hence $T^{-1}(p)$ would separate A contrary to the fact that T is non-separating.

THEOREM 3.2. *If T_2 is non-separating and T_1 is monotone, then T is non-separating.*

Proof. If T were not non-separating, then there would exist a point p of B so that $A - T^{-1}(p) = A_1 + A_2$ would be a separation. Hence, since T_1 is monotone, we would likewise have the separation[6]

$$A' - T_1T^{-1}(p) = T_1(A_1) + T_1(A_2).$$

But $T_1T^{-1}(p) = T_2^{-1}(p)$. Hence $T_2^{-1}(p)$ would separate A contrary to the fact that T_2 is non-separating.

Some simple examples will show us that, if T_1 is non-separating, then T need not be non-separating even though T_2 be both monotone and non-separating. Also T and T_2 may both be non-separating and yet T_1 need not be non-separating even though it be monotone.

In view of G. T. Whyburn's factor theorem for continuous transformations,[7] we now have the following factor theorem for non-separating transformations.

THEOREM 3.3. *If A is a compact continuum and $T(A) = B$ is non-separating, there exist transformations T_1 and T_2 such that $T(x) \equiv T_2T_1(x)$, for $x \in A$, where*

[6] See the proof of a similar theorem of G. T. Whyburn's for non-alternating transformations, loc. cit., p. 296.

[7] Loc. cit., p. 297.

T_1 is monotone, dim $T_2^{-1}(b) = 0$, *for each $b \epsilon B$, and both T_1 and T_2 are non-separating.*

Proof. Let us define the transformations $T_1(A) = A'$ and $T_2(A') = B$ as in Whyburn's factor theorem. Then T_1 is monotone and, for each $b \epsilon B$, dim $T_2^{-1}(b) = 0$. Furthermore T_2 is non-separating by Theorem 3.1. Hence we need only prove that T_1 is non-separating. Now by definition, $T_1(A) = A'$ is such that, for each point p of A', $T_1^{-1}(p)$ is a component of the set $T^{-1}[T_2(p)]$. If we assume that T_1 is not non-separating, there exists some point p of A' such that $T_1^{-1}(p)$ separates A, say $A - T_1^{-1}(p) = A_1 + A_2$. Now $A_1 + T_1^{-1}(p)$ and $A_2 + T_1^{-1}(p)$ are each connected sets, since A and $T_1^{-1}(p)$ are each connected. Hence, if we let $T_2(p) = b \epsilon B$, $T^{-1}(b)$ does not contain A_i, for $i = 1$ or 2. For, if $T^{-1}(b) \supset A_i$, then $T_1^{-1}(p)$ would not be a component of $T^{-1}(b)$, since $A_i + T_1^{-1}(p)$ is connected. Therefore

$$A - T^{-1}(b) = [A_1 - A_1 \cdot T^{-1}(b)] + [A_2 - A_2 \cdot T^{-1}(b)]$$

mutually separated, contrary to the fact that T is non-separating. Hence T_1 is non-separating.

4. **Locally connected continua.** If the continua A and B in Theorem 2.1 are locally connected, that theorem may be rewritten to give us the following

THEOREM 4.1. *If A and B are compact locally connected continua, a necessary and sufficient condition in order that $T(A) = B$ be non-separating is that T be non-alternating and B be cyclically connected.*

It obviously follows from this theorem that the property of being a compact cyclically connected continuum is invariant under non-separating transformations.

With the use of Theorem 2.2 we will now obtain another necessary and sufficient condition for T to be non-separating when A is locally connected.

THEOREM 4.2. *If A and B are compact locally connected continua, a necessary and sufficient condition in order that $T(A) = B$ be non-separating is that, for any $p \epsilon A$, there is at most one component C of $A - p$ on which T is not constant, and $T^{-1}T(p) \cdot C$ does not separate C in A.*

Proof. Take any point p of A. Let C be a component of $A - p$ on which T is not constant. We first prove the necessity. If $C = A - p$, the conclusion is immediate, since T is non-separating. If $C \neq A - p$, we have the separation $A - p = C + (A - p - C)$. Now, by Theorem 2.2, $A - p - C \subset T^{-1}T(p)$, since $C \not\subset T^{-1}T(p)$. Furthermore the set $C - C \cdot T^{-1}T(p)$ is connected. Next we prove the sufficiency. If $C = A - p$, then $C - C \cdot T^{-1}T(p) = A - T^{-1}T(p)$ is connected and hence T is non-separating, since $T(p)$ is any point of B. If $C \neq A - p$, then $T(A - p - C) = T(p)$, by hypothesis and because of the continuity of T. Hence $A - T^{-1}T(p) = C - C \cdot T^{-1}T(p)$, which is connected by hypothesis. Therefore T is non-separating also in this case.

THEOREM 4.3. *If A and B are compact locally connected continua and*

$T(A) = B$ is non-separating, then (I) *there exists a unique true cyclic element E_a of A such that $T(E_a) = B$;* (II) *$T(E_a) = B$ is non-separating; and* (III) *there exists a monotone retracting transformation $W(A) = E_a$ which is non-separating and such that $T(x) \equiv TW(x)$ on A.*

Proof. Since T is non-separating, it is non-alternating. Furthermore B is cyclically connected, by Theorem 4.1. Therefore G. T. Whyburn's Theorems (3.5) and (3.3)[8] establish the existence of the true cyclic element E_a such that $T(E_a) = B$, and the existence of the monotone retracting transformation $W(A) = E_a$ such that $T(x) \equiv TW(x)$ on A. Therefore we need only show that E_a is unique, that T is non-separating on E_a, and that W is non-separating. The uniqueness of E_a follows immediately from Theorem 4.2 since, as a result of that theorem, T must be constant on every component of $A - E_a$, because it is not constant on E_a. If $T(E_a) = B$ were not non-separating, then, for some $b \in B$, $T^{-1}(b) \cdot E_a$ would separate E_a. Then $T^{-1}(b)$ would separate A, since E_a is an A-set.[9] That $W(A) = E_a$ is non-separating follows at once from the fact that, for any point p of E_a, $W^{-1}(p)$ is the point p or the connected set $\bar{K}$ where K is the component of $A - E_a$ whose boundary is p, if such a component exists. In either case $A - W^{-1}(p)$ is connected.

As a result of this theorem we see that the study of the application of non-separating transformations to compact locally connected continua reduces to a study of the application of such transformations to compact cyclically connected continua.

5. **Special curves and surfaces.** In this section we will study the effect of the application of non-separating transformations to several kinds of special curves and surfaces.

(5.1) As a consequence of Theorem 4.3 we have that *the image of any dendrite, and hence of any arc, under any non-separating transformation is a single point.*

(5.2) *Any non-separating transformation $T(A) = B$ defined over a simple closed curve A is monotone. Therefore the property of being a simple closed curve is invariant under non-separating transformations.*

Proof. If T were not monotone, then there would exist some $b \in B$ so that $T^{-1}(b) = C_1 + C_2$, mutually separated. Take a point p_1 of C_1 and a point p_2 of C_2. Since A is a simple closed curve, we have a separation $A - (p_1 + p_2) = A_1 + A_2$. Now $C_1 \not\supset A_i$, for $i = 1$ or 2. For if $C_1 \supset A_i$, then $\bar{C}_1 \supset p_2$ and hence $\bar{C}_1 \cdot C_2 \neq 0$. This contradicts the fact that C_1 and C_2 are mutually separated. Likewise $C_2 \not\supset A_i$, for $i = 1$ or 2. Therefore $C_1 + C_2$ does not contain either A_i, since the A_i are each connected. Hence we have the separation

$$A - (C_1 + C_2) = A - T^{-1}(b) = [A_1 - (C_1 + C_2) \cdot A_1] + [A_2 - (C_1 + C_2) \cdot A_2],$$

[8] Loc. cit., p. 299.

[9] For properties of true cyclic elements and A-sets see C. Kuratowski and G. T. Whyburn, Fundamenta Mathematicae, vol. 16 (1930), pp. 305–331.

contrary to the fact that T is non-separating. Accordingly $T(A) = B$ is monotone and hence B is a simple closed curve, since the property of being a simple closed curve is invariant under monotone transformations.[10]

It is obvious that any monotone transformation defined over a simple closed curve is non-separating.

(5.3) *If A is a boundary curve and $T(A) = B$ is non-separating, then B is a simple closed curve.*

This result is a direct consequence of Theorems 4.3 and 5.2.

In view of the fact that the true cyclic element E_a of A such that $T(E_a) = B$ is a simple closed curve when A is a boundary curve, these two theorems also show us that *any non-separating transformation defined over a boundary curve is monotone,* since $T(E_a) = B$ and $W(A) = E_a$ are both monotone and therefore $T(A) = TW(A) = B$ is monotone.[11]

(5.4) *Let* $A = \sum_{i=1}^{n} a_1 x_i a_2$, *where $a_1 x_i a_2$ are arcs for all i, such that*

$$a_1 x_j a_2 \cdot a_1 x_k a_2 = a_1 + a_2,$$

for $j \neq k$. Then, if $T(A) = B$ is monotone and non-separating, B is a θ-curve or a simple closed curve according as $T(a_1) \neq T(a_2)$ or $T(a_1) = T(a_2)$.

Proof. If $T(a_1) = b_1 \neq b_2 = T(a_2)$, then T is monotone on each arc $a_1 x_i a_2$. Hence $T(a_1 x_i a_2)$, for each i, is an arc joining b_1 and b_2 in B.[12] Moreover any two of the arcs $T(a_1 x_i a_2)$ intersect only in the points b_1 and b_2, since, for any $b \in B - (b_1 + b_2)$, $T^{-1}(b) \subset a_1 x_i a_2$, for some i, because T is monotone and $T(a_1) \neq T(a_2)$. Therefore B is a θ-curve.

If $T(a_1) = T(a_2) = b$, $T^{-1}(b)$ contains all but one, say $a_1 x_k a_2$, of the arcs $a_1 x_i a_2$, since T is monotone and non-separating. Now for any point p of $B - b$, $T^{-1}(b)$ is connected and is contained in $a_1 x_k a_2$. Take any two points y and z of B. We then have the separation $A - T^{-1}(y) - T^{-1}(z) = A_1 + A_2$, where $\bar{A}_1$ is an arc of $a_1 x_k a_2$ which joins a point of $T^{-1}(y)$ to a point of $T^{-1}(z)$ and does not intersect either of these sets in any other points, and

$$A_2 = A - T^{-1}(y) - T^{-1}(z) - A_1.$$

Therefore, since T is monotone, we have the separation

$$B - y - z = T(A_1) + T(A_2),$$

and hence B is a simple closed curve.

(5.41) If we remove the restriction that T be monotone in the above statement, and if $T(a_1) \neq T(a_2)$, $T^{-1}[T(a_1)]$ and $T^{-1}[T(a_2)]$ are each connected sets. Furthermore no set $T^{-1}(b)$ can intersect any arc $a_1 x_i a_2$ in more than one component, and no set $T^{-1}(b)$ which does not contain a_1 or a_2 can intersect more than $n - 1$ of the arcs $a_1 x_i a_2$, since T is non-separating. Hence every set $T^{-1}(b)$ contains at most $n - 1$ components. If $T(a_1) = T(a_2) = b$, it follows

[10] R. L. Moore, loc. cit.

[11] G. T. Whyburn, loc. cit., p. 297, Theorem (2.2).

[12] R. L. Moore, loc. cit.

from the proof of (5.4) that T is also monotone and hence that B is a simple closed curve.

(5.42) If we do not require that T be non-separating in (5.4) but keep the condition that it be monotone, and if $T(a_1) \neq T(a_2)$, it follows from the proof of (5.4) that T is non-separating and hence that B is a θ-curve. If $T(a_1) = T(a_2) = b$, then B is a simple closed curve or $B = \sum_{i=1}^{m} B_i$, where each B_i is a simple closed curve; $2 < m \leqq n - 2$; and $B_k \cdot B_j = b$, for $k \neq j$, by an argument similar to that used in (5.4).

(5.43) In view of Theorem 4.3 and the results of (5.4) we have that *if M is a compact locally connected continuum, each true cyclic element of which is a curve of the same type as A in* (5.4), *and if $T(M) = N$ is monotone and non-separating, then N is a θ-curve or a simple closed curve.*

(5.44) *If A is a θ-curve, or more generally, if A is a compact locally connected continuum each true cyclic element of which is a θ-curve, then, if $T(A) = B$ is monotone and non-separating, B is a simple closed curve or a θ-curve.*

This result follows immediately from (5.4) and (5.43), since a θ-curve is the sum of three arcs $a_1 x_i a_2$ any two of which have only a_1 and a_2 in common.

(5.5) In G. T. Whyburn's Theorem (3.7),[13] if we require that his non-alternating transformation $T(A) = B$ be non-separating, we have *If A is a compact locally connected continuum which is unicoherent and $T(A) = B$ is non-separating, then B is a cantorian manifold of dimension $\geqq 2$.*

(5.6) Let A be a topological sphere. We may now state R. L. Moore's well known theorem[14] that the hyperspace of any upper semi-continuous decomposition of a topological sphere A into continua not separating A is a topological sphere, in the following way: *If A is a topological sphere and $T(A) = B$ is monotone and non-separating, then B is a topological sphere.*

(5.61) If we remove the restriction in (5.6) that T be non-separating, then another theorem of Moore's[15] tells us that B is a cactoid.

(5.62) If we let T be non-separating in (5.6) but do not require that it be monotone, then it follows from (5.5) that B is a cantorian manifold of dimension $\geqq 2$.

(5.7) Now let A be any cactoid. R. L. Moore has shown[16] that monotone transformations carry cactoids into cactoids. Now we have that *if A is a cactoid and $T(A) = B$ is monotone and non-separating, then B is a topological sphere.* For by Theorem 4.3, there exists a true cyclic element E_a of A so that $T(E_a) = B$ and $T(E_a) = B$ is non-separating. Furthermore $T(E_a) = B$ is monotone, since $T^{-1}(b) \cdot E_a$ is connected, for every $b \in B$, because the common part of a connected set and an A-set is connected, if they intersect. Therefore, by the theorem of Moore's stated in (5.6), B is a topological sphere.

Colgate University.

[13] Loc. cit., p. 300.

[14] Loc. cit.

[15] Monatshefte für Mathematik und Physik, vol. 36 (1929), pp. 81–88. Also see G. T. Whyburn, loc. cit., p. 300.

[16] Loc. cit.

nple

the
roof
) =
is a
nent

is a
urve
ing,

cted
B is

the

ter-
com-
non-

well
osi-
gical
B is

hen

t be
nen-

one
is a
nical
that
B is
mon
fore,

. T.